Limit Laws in Transaction-Level Asset Price Models

Alexander Aue* Lajos Horváth† Clifford Hurvich‡
Philippe Soulier§

January 19, 2013

Abstract

We consider pure-jump transaction-level models for asset prices in continuous time, driven by point processes. In a bivariate model that admits cointegration, we allow for time deformations to account for such effects as intraday seasonal patterns in volatility, and non-trading periods that may be different for the two assets. We also allow for asymmetries (leverage effects). We obtain the asymptotic distribution of the log-price process. We also obtain the asymptotic distribution of the ordinary least-squares estimator of the cointegrating parameter based on data sampled from an equally-spaced discretization of calendar time, in the case of weak fractional cointegration. For this same case, we obtain the asymptotic distribution for a tapered estimator under more general assumptions. In the strong fractional cointegration case, we obtain the limiting distribution of a continuously-averaged tapered estimator as well as other estimators of the cointegrating parameter, and find that the rate of convergence can be affected by properties of intertrade durations. In particular, the persistence of durations (hence of volatility) can affect the degree of cointegration. We also obtain the rate of convergence of several estimators of the cointegrating parameter in the standard cointegration case. Finally, we consider the properties of the ordinary least squares estimator of the regression parameter in a spurious regression, i.e., in the absence of cointegration.

*Department of Statistics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA, email: `alexaue@wald.ucdavis.edu`

†Department of Mathematics, University of Utah, 155 South 1440 East, Salt Lake City, UT 84112, USA, email: `horvath@math.utah.edu`

‡Stern School of Business, New York University, Henry Kaufman Management Center, 44 West Fourth Street, New York, NY 10012, USA, email: `churvich@stern.nyu.edu`

§Université Paris X, 200 avenue de la République, 92001 Nanterre cedex, France, email: `philippe.soulier@u-paris10.fr`

1 Introduction

The increasingly widespread availability of transaction-level financial price data motivates the development of models to describe such data, as well as theory for widely-used statistics of interest under the assumption of a given transaction-level generating mechanism. We focus here on a bivariate pure-jump model in continuous time for log prices proposed by Hurvich and Wang (2009, 2010) which yields fractional or standard cointegration. The motivation for using a pure-jump model is that observed price series are step functions, since no change is possible in observed prices during time periods when there are no transactions. Examples of data sets that would fit into the framework of this model include: buy prices and sell prices of a single stock; prices of two different stocks within the same industry; stock and option prices of a given company; option prices on a given stock with different degrees of maturity or moneyness; corporate bond prices at different maturities for a given company; Treasury bond prices at different maturities.

Though our paper is not entirely focused on the case of fractional cointegration, we present here some evidence that this case may arise in practice in financial econometrics. We considered option and underlying best-available bid prices for 69 different options on IBM at 390 one-minute intervals from 9:30 AM to 4 PM on May 31, 2007. Using a log-periodogram estimator based on $390^{0.5}$ frequencies, we found that the logs of the original series had estimated memory parameters close to 1, while the residuals from the OLS regression of the log stock price on the log option price had estimated memory parameters that were typically less than 1. Specifically, of the 69 estimated memory parameters based on these residuals, the values ranged from 0.05 to 1.14 with a mean of 0.55 and a standard deviation of 0.28, with 30 of these estimates lying between 0.5 and 1, while 32 were between 0 and 0.5. Thus, there is evidence for cointegration in most of the series studied, and often the evidence points towards fractional rather than standard cointegration. Furthermore, the OLS estimate of the cointegrating parameter (assuming that cointegration exists) ranged from -0.21 to 0.39, with a mean of 0.04 and a standard deviation of 0.13. This provides evidence that the cointegrating parameter is in general not equal to one in the present context, so it is of interest to study properties of estimates of this parameter.

Two basic questions that we address in this paper are the asymptotic distribution of the log prices as time $t \to \infty$, and of the usual OLS estimator of the cointegrating parameter based on n observations of the log prices at equally-spaced time intervals as $n \to \infty$. Most of the existing methods for deriving such limit laws (see Robinson and Marinucci (2001)) cannot be applied here because the continuous-time log-price series are not diffusions and because the discretized log-price series are not linear in either an *iid* sequence, a martingale difference sequence or a strong mixing sequence. Nevertheless, it is of interest to know whether and under what conditions the existing limit laws, based, say, on linearity assumptions in discrete time, continue to hold under a transaction-level generating mechanism.

In the model of Hurvich and Wang (2010, 2009) the price process in continuous time is

specified by a counting process giving the cumulative number of transactions up to time t, together with the process of changes in log price at the transaction times. This structure corresponds to the fact that most transaction-level data consists of a time stamp giving the transaction time as well as a price at that time. In such a setting, another observable quantity of interest is the *durations*, i.e., the waiting times between successive transactions of a given asset. There is a growing literature on univariate models for durations, including the seminal paper of Engle and Russell (1998) on the autoregressive conditional duration models (ACD), as well as Bauwens and Veredas (2004) on the stochastic duration model (SCD), and Deo et al. (2010) on the long-memory stochastic duration model (LMSD).

Deo et al. (2009) showed that, subject to regularity conditions, if partial sums of centered durations, scaled by $n^{-(d+1/2)}$ with $d \in [0, 1/2)$, satisfy a functional central limit theorem then the counting process $N(t)$ has long or short memory (for $d > 0$, $d = 0$, respectively) in the sense that $\mathrm{Var}N(t) \sim Ct^{2d+1}$ as $t \to \infty$ (with $C > 0$), and they gave conditions under which this scaling would lead to long memory in volatility. In particular, LMSD durations with $d > 0$ lead to long memory in volatility. The latter property has been widely observed in the econometrics literature, while evidence for long memory in durations was found in Deo et al. (2010).

Hurvich and Wang (2010, 2009) did not derive limit laws for the log price series or the OLS estimator of the cointegrating parameter, but focused instead on properties of variances and covariances for log price series and returns, and on lower bounds on the rate of convergence for the OLS estimator.

In this paper, for a slightly modified version of the model of Hurvich and Wang (2010, 2009), but under assumptions that are more general than theirs, we obtain the limit law for the log prices, and for the OLS and tapered estimators of the cointegrating parameter. In our result on the limit law for log prices, Theorem 3.1, we allow for a stochastic time-varying intensity function in the counting processes. This allows for such effects as dynamic intraday seasonality in volatility (as observed, for example, in Deo et al. (2006), as well as fixed nontrading intervals such as holidays and overnight periods. We also allow in most of our results for asymmetries (leverage effects), and show that this opens up the possibility that long memory in durations may affect the rate of convergence of estimators of the cointegrating parameter. This raises some heretofore unrecognized ambiguities in the choice of a definition of standard cointegration. Finally, we consider the properties of the ordinary least squares estimator in a spurious regression, i.e., in the absence of cointegration.

The remainder of this paper is organized as follows. In Section 2 we write the model for the log price series and state our assumptions on the counting process, the time-deformation functions, and the return shocks. In Section 3, we provide our main results on: the long-run behavior of the log-price process (Subsection 3.1), the OLS estimator for the cointegrating parameter under weak fractional, strong fractional and standard cointegration (Subsection 3.2), a tapered estimator under weak fractional, strong fractional and standard cointegration (Subsection 3.3), a continuously-averaged tapered estimator under strong

fractional and standard cointegration (Subsection 3.4) and the ordinary least squares estimator in the spurious regression case (Subsection 3.5). Section 4 provides proofs.

2 Transaction level model

As in Hurvich and Wang (2010, 2009), we consider a bivariate pure-jump transaction-level price model that enables cointegration. We define the log-price process $y = (y_1, y_2) = (y(t)\colon t \geq 0)$ by

$$y_1(t) = \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \theta \sum_{k=1}^{N_2(t_{1,N_1(t)})} e_{2,k}, \tag{2.1}$$

$$y_2(t) = \sum_{k=1}^{N_2(t)} (e_{2,k} + \eta_{2,k}) + \theta^{-1} \sum_{k=1}^{N_1(t_{2,N_2(t)})} e_{1,k}, \tag{2.2}$$

where for $i = 1, 2$, $N_i(\cdot)$ are counting processes on the real line (see Daley and Vere-Jones (2003, page 43)) such that, for $t \geq 0$, $N_i(t) := N_i(0, t]$ gives the total number of transactions of Asset i in $(0, t]$, and $t_{i,k}$ is the clock time (calendar time) for the kth transaction of Asset i, with $\cdots t_{i,-1} \leq t_{i,0} \leq 0 < t_{i,1} \leq t_{i,2} \cdots$. The quantity $N_2(t_{1,N_1(t)})$ denotes the number of transactions of Asset 2 between time 0 and the time $t_{1,N_1(t)}$ of the most recent transaction of Asset 1, with an analogous interpretation for $N_1(t_{2,N_2(t)})$. The efficient shock sequences $\{e_{i,k}\}_{k=1}^{\infty}$ model the permanent component and the microstructure noise sequences $\{\eta_{i,k}\}_{k=1}^{\infty}$ model the transitory component of the log-price process. Efficient shock spillover effects are weighted by θ and θ^{-1}, thus yielding cointegration with cointegrating parameter θ, assumed nonzero. A detailed economic justification for this model, derivation of a common-components representation, as well as a comparison with certain discrete-time models, is given in Hurvich and Wang (2010, 2009).

Following Daley and Vere-Jones (2003, page 47), a point process is said to be *simple* if the probability is zero that there exists a time t at which more than one event occurs. We do not assume that the counting processes are simple. Thus we allow for the possibility that several transactions may occur at the same time. The transaction times $t_{i,k}$ are related to the point process by the following duality.

$$N_i(t) = k \Leftrightarrow t_{i,k} \leq t < t_{i,k+1} \; .$$

The durations are then defined for $k \geq 1$ by

$$\tau_{i,1} = t_{i,1} \; , \; \tau_{i,k} = t_{i,k} - t_{i,k-1} \; .$$

If the process is simple, then $N_i(t_{i,k}) = k$. Otherwise, it only holds that $N_i(t_{i,k}) \geq k$. We need the following ergodicity-type assumptions.

Assumption 2.1. *The sequences* $\{t_{i,k}\}$ *are nondecreasing and there exists* $\lambda_i \in (0, \infty)$ *such that*

$$t_{i,k}/k \xrightarrow{P} 1/\lambda_i \ . \tag{2.3}$$

When the counting processes are simple, this is equivalent to $N_i(t)/t \xrightarrow{P} \lambda_i$. Since we do not assume simplicity, we must introduce an additional assumption.

Assumption 2.2. $N_i(t)/t \xrightarrow{P} \lambda_i$.

If the counting processes are defined from stationary ergodic durations, then Assumption 2.1 holds. If the couting processes are moreover simple, then Assumption 2.2 also holds. Thwo such examples where both Assumptions 2.1 and 2.2 hold are LMSD processes and ACD processes, under conditions specified presently. Here and elsewhere, we omit the i subscript when this does not cause confusion.

- For the LMSD process, suppose that $\tau_k = \epsilon_k e^{\sigma Y_k}$ where σ is a positive constant, $\{\epsilon_k\}$ is an *iid* sequence of almost surely positive random variables with finite mean and $\{Y_k\}$ is a stationary standard Gaussian process, independent of $\{\epsilon_k\}$, such that $\lim_{k\to\infty} \operatorname{cov}(Y_0, Y_k) = 0$. It follows from the latter assumption and Guassianity that the process $\{Y_k\}$ is ergodic Ibragimov and Rozanov (1978), hence so is $\{\tau_k\}$. Thus Assumption 2.1 holds with $\lambda^{-1} = \mathbb{E}[\epsilon_1]e^{\sigma^2/2}$ and since the durations are almost surely positive, so does Assumption 2.2. Henceforth, we will further assume that either $\sum_{k=1}^{\infty} |\operatorname{cov}(Y_0, Y_k)| < \infty$ or $\operatorname{cov}(Y_0, Y_k) \sim ck^{2H_\tau - 2}$ for some $c > 0$ and $H_\tau \in (1/2, 1)$.
- The ACD model proposed in Engle and Russell (1998) is

$$\tau_k = \psi_k \epsilon_k, \qquad \psi_k = \omega + \alpha\tau_{k-1} + \beta\psi_{k-1}, \qquad k \in \mathbb{Z}, \tag{2.4}$$

  where $\omega > 0$ and $\alpha, \beta \geq 0$, $\{\epsilon_k\}_{k=-\infty}^{\infty}$ is an *iid* sequence with $\epsilon_0 > 0$ almost surely, and $E[\epsilon_0] = 1$. The sequence $\{\tau_k\}_{k=-\infty}^{\infty}$ is then strictly stationary provided such a solution to (2.4) exists. From Grimmett and Stirzaker (2001, Section 9.5), it follows that stationary ACD durations are ergodic if $E[\tau_0] < \infty$. The strictly stationary solution is then determined by $\psi_k = \omega \sum_{j=1}^{\infty} \prod_{i=1}^{j-1} (\alpha\epsilon_{k-i} + \beta)$ and exists if $E[\ln(\alpha\epsilon_0 + \beta)] < 0$, following arguments given in Bougerol and Picard (1992). As in Aue et al. (2006), one can derive that now $E[\tau_0]$ is finite if $\alpha + \beta < 1$ which, on account of Jensen's inequality, is also sufficient for $E[\ln(\alpha\epsilon_0 + \beta)] < 0$ to hold. Thus, $\{\tau_k\}$ is ergodic as long as $\alpha + \beta < 1$. In that case, Assumptions 2.1 and 2.2 hold with $\lambda = (\mathbb{E}[\tau_1])^{-1} = (1 - \alpha - \beta)/\omega$.

We now give an example (which is allowed as a special case of our theorems) of how time deformations can be used to obtain a nonstationary, possibly non simple point process from a stationary simple one. Let $\tilde{N}_i(\cdot)$ be a simple, stationary and ergodic counting process on

$\mathbb{R}$ with intensity $\tilde{\lambda}_i \in (0, \infty)$ and let f_i be deterministic or random functions such that f_i is nondecreasing and has càdlàg paths with probability one. For $i = 1, 2$, define then

$$N_i(t) = \tilde{N}_i(f_i(t)) \ .$$

If the functions f_i are random, we assume moreover that they are independent of the counting processes $\tilde{N}_i$.

The functions f_i are used to speed up or slow down the trading clock. To incorporate dynamic intraday seasonality in volatility, the same time deformation can be used in each trading period (of length, say, T), assuming that $f_i(t)$ has a periodic derivative (with period T and with probability one), for example, $f_i(t) = t + .5 \sin(2\pi t/T)$. Fixed nontrading intervals, say, $t \in [T_1, T_2)$, could be accommodated by taking $f_i(t) = f_i(T_1)$ for $t \in [T_1, T_2)$ so that $f_i(t)$ remains constant for t in this interval, and then taking $f_i(T_2) > f_i(T_1)$ so that $f_i(t)$ jumps upward when trading resumes at time T_2. The jump allows for the possibility of one or more transactions at time T_2, potentially reflecting information from other markets or assets that did trade in the period $[T_1, T_2)$.

If the values of series i are only recorded at specific time points (e.g., quarterly in the case of certain macroeconomic series) this could be handled by taking $f_i(t)$ to be a pure-jump function. This provides scope for considering two (or more) series, some of which are observed continuously, others at specific times, though not necessarily contemporaneously. In future work, we hope to explore this scenario in detail, and its possible connections with the MIDAS methodology, see Ghysels et al. (2006).

The use of the time-varying intensity function f_i renders the counting process N_i non-stationary. Since it is possible that f_i has (upward) jumps, the N_i may also not be simple even though the $\tilde{N}_i$ are simple. We now show, however, that Assumptions 2.1 and 2.2 hold under the present assumptions.

Lemma 2.1. *Assume that f is a nondecreasing (random) function such that $t^{-1} f(t) \xrightarrow{P} \gamma \in (0, \infty)$ and*

$$\sup_{t \geq 0} |f(t) - f(t^-)| \leq C$$

with probability one, where $C \in (0, \infty)$ is a deterministic constant. Let $\tilde{N}$ be a stationary ergodic simple point process, so that Assumptions 2.1 and 2.2 hold for some $\lambda > 0$. Let N be the counting process defined by $N(\cdot) = \tilde{N}(f(\cdot))$. Then Assumptions 2.1 and 2.2 hold for N with $\lambda = \tilde{\lambda}\gamma$.

In order to show that our results on estimation of the cointegrating parameter (under weak fractional and standard cointegration) hold in this time deformation framework, we will in Lemma 4.7 make further assumptions on the f_i. These assumptions mathematically embody natural economic constraints, viz. minimum duration of trading and nontrading periods, maximum duration of nontrading periods and non stoppage of trading time during trading periods.

We now state our assumptions on the return shocks.

Assumption 2.3. *The efficient shocks $\{e_{i,k}\}$ are mutually independent i.i.d. sequences with zero mean and variance $\sigma_{i,e}^2$.*

Although many of our results would continue to hold if the *iid* part of Assumption 2.3 were replaced by a weak-dependence assumption, we maintain the *iid* assumption here in keeping with the economic motivation for the model as provided by Hurvich and Wang (2010) that in the absence of the microstructure shocks and in the absence of any dependence of the efficient shocks on the counting processes, each of the log price series would be a martingale with respect to its own past. Since the trades of Asset 1 are not synchronized in calendar time or in transaction time with those of Asset 2, it seems reasonable to assume that the two efficient shock series are mutually independent, as we have done in Assumption 2.3.

The following assumption implies that the microstructure noise does not affect the limiting distribution of the log prices. We use $\Rightarrow$ to denote weak convergence in the space $\mathcal{D}([0,\infty))$ of left-limited, right-continuous (càdlàg) functions, endowed with Skorohod's J_1 topology. See Billingsley (1968) or Whitt (2002) for details about weak convergence in $\mathcal{D}([0,\infty))$. Whenever the limiting process is continuous, this topology can be replaced by the topology of uniform convergence on compact sets.

Assumption 2.4 (Microstructure Noise)**.** *The microstructure noise sequences $\{\eta_{i,k}\}$ satisfy $n^{-1/2}\sum_{k=1}^{[n\cdot]}\eta_{i,k} \Rightarrow 0$.*

Dependence between the counting processes and return shocks allows for leverage effects (for example, a correlation between a return in one time period and a squared return in a subsequent time period). A transaction-level model yielding a leverage effect was proposed (but justified only with simulations) in Hurvich and Wang (2009). Models where the point process need not be independent of the return shocks were discussed in Prigent (2001) in the context of option pricing with marked point processes.

We do not make any assumption of independence between the counting processes and the microsctucture shocks, unless explicitly noted otherwise. We will, however, assume that the counting processes are independent of the efficient shocks except in Theorem 3.1 and in Section 3.5.

Assumption 2.4 is all we need to assume about the microstructure noise in order to obtain a limit law for the log price series (such as Theorem 3.1 below). However, in order to discuss properties of estimators of the cointegrating parameter it is necessary to make more specific assumptions on the degree of cointegration. In Hurvich and Wang (2009, 2010), three different cases were considered, according to the strength of the memory of the microstructure noise sequences. These cases were labeled as weak fractional cointegration, strong fractional cointegration and standard cointegration. In the current context, where there may be a dependence between return shocks and counting processes, special care is needed in defining the strong fractional and standard cointegration cases, as long memory in durations may affect the rate of convergence of estimators of the cointegrating parameter

in these cases. On the other hand, we will define weak fractional cointegration essentially as in Hurvich and Wang (2010).

Assumption 2.5. *The shocks* $\{e_{1,k}\}_{k=-\infty}^{\infty}$, $\{e_{2,k}\}_{k=-\infty}^{\infty}$, $\{\eta_{1,k}\}_{k=-\infty}^{\infty}$ *and* $\{\eta_{2,k}\}_{k=-\infty}^{\infty}$ *are mutually independent.*

Mutual independence of the efficient and microstructure shock series of a given asset can be justified on economic grounds, and is often made in the econometric literature for calendar-time models. See, e.g., Barndorff-Nielsen et al. (2008). Mutual independence of the two microstructure series is justified by the lack of synchronization of the trading times of the two assets.

We now discuss the weak fractional cointegration case. For $H \in (0,1)$, let B_H denote the standard fractional Brownian motion (FBM) with Hurst index H, i.e. the zero mean Gaussian process with almost surely continuous sample paths and covariance function

$$\mathrm{cov}(B_H(s), B_H(t)) = \frac{1}{2}\left\{s^{2H} - |t-s|^{2H} + t^{2H}\right\} .$$

For $H = 1/2$, $B_{1/2}$ is the standard Brownian motion.

Assumption 2.6 (Weak Fractional Cointegration)**.** *There exists* $H \in (0,1/2)$ *such that*

$$n^{-H}\sum_{k=1}^{[n\cdot]} \eta_{i,k} \Rightarrow c_i B_H^{(i)}$$

where c_1, c_2 *are nonnegative constants, not both zero and* $B_H^{(1)}$ *and* $B_H^{(2)}$ *are independent standard fractional Brownian motions with common Hurst index* H*.*

Under Assumption 2.5, the independence of all the noise series implies that all the previous convergences hold jointly. The situation where one of the constants c_1 or c_2 is zero could arise naturally if the memory in one of the microstructure noise series is weaker than for the other.

In the case of weak fractional cointegration, we can define the memory parameter of the microstructure noise series as $d_\eta = H - 1/2 \in (-1/2, 0)$, and the degree of fractional cointegration (i.e. the rate of convergence of partial sums of the cointegrating error) is completely determined by d_η. More precisely, in this case the difference between the memory parameters of the series of log prices and the cointegrating error (observed, say, at equally-spaced intervals of calendar time) $y_1(j) - \theta y_2(j)$, is $-d_\eta$. This holds regardless of any dependence between the counting processes and the microstructure shocks.

Next we discuss strong fractional and standard cointegration. We start with the assumption that, for $i = 1, 2$, $\eta_{i,k} = \xi_{i,k} - \xi_{i,k-1}$ where $\{\xi_{i,k}\}$ satisfy $\sup_k \mathbb{E}[\xi_{i,k}^2] < \infty$. It then follows that the cointegrating error at time j is

$$
\begin{aligned}
&y_1(j) - \theta y_2(j) \\
&\quad = \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} - \theta \sum_{k=N_2(t_{1,N_1(j)})+1}^{N_2(j)} e_{2,k} + \xi_{1,N_1(j)} - \xi_{1,0} - \theta(\xi_{2,N_2(j)} - \xi_{2,0}). \quad (2.5)
\end{aligned}
$$

Under the assumptions we will make in this paper, and also under the assumptions made in Hurvich and Wang (2010, 2009), the first two terms on the righthand side of (2.5) are weakly dependent, so the degree of cointegration is determined by the rate of convergence of partial sums of $\xi_{i,N_i(j)}$.

Thus we will need to study the sequence $\xi_{i,N_i(j)}$. We do not assume that the microstructure shocks are independent of the counting processes. Thus, even if the microstructure shocks have zero mean, it may hold that $\mathbb{E}[\xi_{i,N_i(j)}] \neq 0$.

Assumption 2.7 (Strong fractional and standard cointegration)**.** *The microstucture noise sequences $\{\eta_{i,k}\}$ can be expressed as $\eta_{i,k} = \xi_{i,k} - \xi_{i,k-1}$. There exist $H \in [1/2, 1)$, constants μ_1^*, μ_2^* and nonnegative constants c_1, c_2, not both zero, such that*

$$
n^{-H} \sum_{k=1}^{[n\cdot]} \{\xi_{i,N_i(j)} - \mu_i^*\} \Rightarrow c_i B_H^{(i)}
$$

where $B_H^{(1)}$ and $B_H^{(2)}$ are independent fractional Brownian motions with Hurst index H.

The case $H > 1/2$ corresponds to strong fractional cointegration whereas the case $H = 1/2$ corresponds to standard cointegration.

It might be hard to verify Assumption 2.7 unless the durations are integer valued. Since commonly-used duration models do not have integer-valued durations, we will introduce a modification of the estimators which involves integrals instead of sums, thus allowing us to avoid this restriction. This change requires a corresponding modification of Assumption 2.7.

Assumption 2.8 (Strong fractional and standard cointegration)**.** *The microstucture noise sequences $\{\eta_{i,k}\}$ can be expressed as $\eta_{i,k} = \xi_{i,k} - \xi_{i,k-1}$. There exist $H \in [1/2, 1)$, constants μ_1^*, μ_2^* and nonnegative constants c_1, c_2, not both zero, such that*

$$
n^{-H} \int_0^{n\cdot} \{\xi_{i,N_i(s)} - \mu_i^*\}\, \mathrm{d}s \Rightarrow c_i B_H^{(i)} \ .
$$

In their strong fractional cointegration case, Hurvich and Wang (2010) assumed, for $d_\eta \in (-1, -1/2)$, that $\mathrm{cov}(\xi_{i,k}, \xi_{i,k+j}) \sim K j^{2d_\xi - 1}$ as $j \to \infty$ where $K > 0$ and $d_\xi = d_\eta + 1 \in (0, 1/2)$. They then showed (in their Lemma 3), under the assumptions made there, that $\mathrm{cov}(\xi_{i,N_i(k)}, \xi_{i,N_i(k+j)}) \sim K' j^{2d_\xi - 1}$ as $j \to \infty$ where $K' > 0$, so that the degree of fractional cointegration was completely determined by the rate of decay of $\mathrm{cov}(\xi_{i,k}, \xi_{i,k+j})$.

However, the proof of this result relied on the assumption that the microstructure shocks are independent of the counting processes, an assumption which we do not make here.

We next provide an example showing that under dependence between the microstructure shocks and the counting processes, it is possible for $\{\xi_{i,k}\}$ to be weakly dependent, and yet the rate of convergence of suitably normalzed integrals of the process $(\xi_{i,N_i(t)} : t \geq 0)$ is determined by the degree of long memory in durations. Suppressing the i subscript, we have the following lemma.

Lemma 2.2. *Suppose that $\{\tau_k\}$ is given by the LMSD model $\tau_k = \epsilon_k \mathrm{e}^{Y_k}$, $\{\epsilon_k\}$ are i.i.d. standard exponential, independent of the stationary standard Gaussian series $\{Y_k\}$, which satisfies $\mathrm{cov}(Y_0, Y_n) \sim cn^{2H_\tau - 2}$ where $c > 0$, and $H_\tau \in (1/2, 3/4)$. Define $\xi_k = Y_{k+1}^2 - 1$. Then the autocovariance function of $\{\xi_k\}$ is summable and there exists $c' > 0$ such that*

$$n^{-1/2} \sum_{k=1}^{[n\cdot]} \xi_k \Rightarrow c' B \; . \tag{2.6}$$

Nevertheless, the randomly-indexed continuous-time process $\xi_{N(t)}$ has long memory in the sense that there exists a constant μ^*, as well as a constant c'', such that

$$n^{-H_\tau} \int_0^{n\cdot} \{\xi_{N(s)} - \mu^*\} \, \mathrm{d}s \Rightarrow c'' B_{H_\tau} \; . \tag{2.7}$$

Lemma 2.2 shows that long memory in durations can induce the same degree of long memory in the cointegrating error (2.5) in calendar-time, even though the microstructure shocks, which are the source of the cointegration, have short memory as a sequence in transaction time. In Lemma 2.2, this phenomenon was achieved by imposing a particular functional relationship between the (zero mean) microstructure noise and the persistent component of the durations, $\xi_k = Y_{k+1}^2 - 1$. This relationship implies a leverage effect, since $\mathrm{corr}(\xi_k, \tau_{k+1}) = 1/\sqrt{2(e-1)} \approx .539 > 0$. In other words, a strongly negative microstructure shock to the return leads to a tendency of the next observed duration, as well as subsequent durations, to be shorter than average. Such a string of short durations increases the volatility, e.g., the expectation of squared calendar-time returns, as shown, for example, under a particular return model in Deo et al. (2009). Furthermore, evidence that stock market intertrade durations have long memory was provided in Deo et al. (2010).

In the absence of dependence between the counting processes and microstructure noise series, in both cases of strong fractional and standard cointegration, the memory of durations cannot affect the memory of the cointegrating error. See Lemma 4.9 for strong fractional cointegration and Lemma 4.10 for standard cointegration.

3 Main results

3.1 The long-run behavior of the bivariate log-price process

With the assumptions made in Section 2, the long-run behavior of the bivariate process $y = (y_1, y_2)$ can be determined. The following theorem shows that the log-prices are approximately integrated. Even though independence is assumed between the various shock series, the log-price process $y = (y(t)\colon t \geq 0)$ exhibits a nontrivial variance-covariance structure which is determined by a complex interplay of the model parameters.

Theorem 3.1. *Under Assumptions 2.1, 2.2, 2.3, 2.4, $n^{-1/2}(y_1(n\cdot), y_2(n\cdot)) \Rightarrow \mathbb{B}$, where*

$$\mathbb{B} = \left(\sigma_{1,e}\sqrt{\lambda_1}B_1 + \theta\sigma_{2,e}\sqrt{\lambda_2}B_2\,,\ \theta^{-1}\sigma_{1,e}\sqrt{\lambda_1}B_1 + \sigma_{2,e}\sqrt{\lambda_2}B_2\right)\,. \tag{3.1}$$

and B_1 and B_2 are independent standard Brownian motions.

In Theorem 3.1, we have not assumed that the counting processes are independent of either the efficient shocks or the microstructure shocks.

Hurvich and Wang (2010, 2009) have in their Theorem 1 computed the long-run variances of $y_1(t)$ and $y_2(t)$ which are given as $(\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2)t$ and $(\sigma_{1,e}^2\lambda_1/\theta^2 + \sigma_{2,e}^2\lambda_2)t$, respectively. Our theorem yields the variances as well as the covariances in the limiting distribution of $(t^{-1/2}y(t)\colon t \geq 0)$. More importantly, our theorem provides the limiting distribution itself for the (normalized) log-price process y which, in turn, can be used for asymptotic statistical inference. Indeed, most of the subsequent results in this paper use Theorem 3.1 and its proof as a building block. In particular, a slightly generalized version of this theorem directly yields asymptotics for estimators in spurious regressions and therefore can be used to motivate tests for the null hypothesis of no cointegration, as we describe in Section 3.5.

3.2 OLS estimator of the cointegrating parameter

In this section, we derive the asymptotic behavior of the ordinary least-squares estimator (OLS) of the cointegrating parameter θ. To do so, we assume that the log-price series are observed at integer multiples of Δt. We will work here, without loss of generality, with $\Delta t = 1$ in order to keep the notation simple. Then (2.1) and (2.2) become

$$y_1(j) = \sum_{k=1}^{N_1(j)} (e_{1,k} + \eta_{1,k}) + \theta \sum_{k=1}^{N_2(t_{1,N_1(j)})} e_{2,k},$$

$$y_2(j) = \sum_{k=1}^{N_2(j)} (e_{2,k} + \eta_{2,k}) + \theta^{-1} \sum_{k=1}^{N_1(t_{2,N_2(j)})} e_{1,k}.$$

Regressing $y_1(1), \dots, y_1(n)$ on $y_2(1), \dots, y_2(n)$ without intercept, we obtain the OLS estimator of θ as

$$\hat{\theta}_n^{OLS} = \frac{\sum_{j=1}^n y_2(j) y_1(j)}{\sum_{j=1}^n y_2^2(j)}. \tag{3.2}$$

Hurvich and Wang (2010, 2009) have shown in their Theorem 6 (under conditions that are for the most part stronger than the ones we assume here) that $\hat{\theta}_n^{OLS}$ is weakly consistent for θ and obtained a lower bound on the rate of convergence in the case of weak fractional, strong fractional and standard cointegration. The exact limit distributions, however, were not given. We fill in this gap next for weak fractional cointegration.

3.2.1 OLS under weak fractional cointegration

Theorem 3.2. *Let Assumptions 2.1, 2.2, 2.3, 2.5 and 2.6 hold. Assume in addition that the counting processes N_1 and N_2 are mutually independent and independent of the efficient shocks and there exists a constant C such that*

$$\sup_{s \geq 0} \mathbb{E}[t_{i,N_i(s)+1} - s] \leq C\ . \tag{3.3}$$

Then

$$n^{1/2-H}(\hat{\theta}_n^{OLS} - \theta) \xrightarrow{d} \Sigma_1 \frac{\int_0^1 B(t) B_H(t)\, \mathrm{d}t}{\int_0^1 B^2(t)\, \mathrm{d}t}$$

where B is a standard Brownian motion, B_H is a fractional Brownian motion, independent of B, and

$$\Sigma_1^2 = \frac{c_1^2 \lambda_1^{2H} + c_2^2 \lambda_2^{2H}}{\theta^{-2} \lambda_1 \sigma_{1,e}^2 + \lambda_2 \sigma_{2,e}^2}\ .$$

The result in Theorem 3.2 is similar to that obtained in Robinson and Marinucci (2001, Proposition 6.5, formula (6.8)), under their Assumption 6.1, for which a sufficient condition (their formula (6.5)) was verified in Marinucci and Robinson (2000) to hold for weak (but not strong) fractional cointegration in the case where the process is linear with respect to *iid* innovations.

Remark 3.1. For the LMSD model described in the previous section, it is proved in Lemma 4.8 that (3.3) holds under the assumptions of Lemma 2.2 that pertain to the durations. For the ACD model described in the previous section, Lemma 4.6 shows that (3.3) holds as long as $\mathbb{E}[\tau_k^3] < \infty$. Carrasco and Chen (2002, Corollary 6) prove that if $\mathbb{E}[(\beta + \alpha\epsilon_t)^3] < 1$ then $\mathbb{E}[\tau_k^3] < \infty$. For the LMSD or ACD models, or for any duration model such that the corresponding counting process $\tilde{N}$ satisfies (3.3), the time deformed process N also satisfies (3.3), provided the time deformation function f satisfes the assumptions of Lemma 4.7. Therefore Theorem 3.2 holds for LMSD or ACD durations with deformation functions satisfying the assumptions of Lemmas 2.1 and 4.7.

3.2.2 OLS under strong fractional and standard cointegration

We now consider the case where the microstructure noise series $\{\eta_{i,k}\}$ are differences of strongly or weakly dependent processes $\{\xi_{i,k}\}$.

Theorem 3.3. *Let Assumptions 2.1, 2.2, 2.3, 2.5, and 2.7 hold. Assume moreover that*

- *the efficient shocks are i.i.d. Gaussian,*
- *the counting processes N_1 and N_2 are independent of each other and independent of the microstructure noise sequences and of the efficient shocks and there exists a constant C such that*

$$\sup_{t\geq 0} \mathbb{E}[(t_{i,N_i(t)+1} - t)^2] \leq C \ . \tag{3.4}$$

- $\mathbb{E}[\xi_{i,k}] = 0$, $\sup_k \mathbb{E}[\xi_{i,k}^2] < \infty$, *and* $\xi_{i,0} = 0$.

Then,

- *if* $1/2 < H < 1$,

$$n^{3/2-H}(\hat{\theta}_n^{OLS} - \theta) \xrightarrow{d} \sqrt{\frac{c_1^2 + \theta^2 c_2^2}{\theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2}} \, \frac{\int_0^1 B(s)\,\mathrm{d}B_H(s)}{\int_0^1 B^2(s)\,\mathrm{d}s} \ , \tag{3.5}$$

  where B is standard Brownian motion independent of the standard fractional Brownian motion B_H;

- *if* $H = 1/2$, $n(\hat{\theta}_n^{OLS} - \theta) = O_P(1)$.

The rate of convergence obtained in the standard cointegration case improves on the one obtained by Hurvich and Wang (2010).

The assumptions in Theorem 3.3 are quite strong, ruling out leverage effects and providing one motivation for our subsequent consideration of tapered estimators.

Remark 3.2. For the LMSD model described in the previous section, it is proved in Lemma 4.8 that (3.4) holds under the assumptions of Lemma 2.2 pertaining to the durations. Similarly, for the ACD durations, Lemma 4.6 and Carrasco and Chen (2002, Corollary 6) imply that if $\mathbb{E}[(\beta + \alpha\epsilon_1)^5] < \infty$ then (3.4) holds. However, we are unable to show that Assumption 2.7 holds except in certain cases of integer-valued durations. We do not pursue this further here.

3.3 A Tapered Estimator of the Cointegrating parameter

Even in existing discrete-time models for cointegration the OLS estimator lacks any particular optimality properties. Here we consider an estimator based on discrete Fourier transforms of the tapered differences of $y_1(j),\ y_2(j),\ 1 \leq j \leq n$. It was shown in Chen and Hurvich (2003a) that this estimator can have a faster rate of convergence than OLS in certain cases of fractional cointegration. In the weak fractional cointegration case, our limit results for the tapered estimator (Theorem 3.4) are obtained under identical conditions as those assumed in Theorem 3.2 for OLS. However, under strong fractional and standard cointegration, the conditions for our results on the tapered estimator (Theorem 3.5) allow for leverage, unlike the corresponding theorem for OLS.

We introduce all relevant notation using a generic time series $\{x_j\}_{j=-\infty}^{\infty}$. Let $h\colon I \to \mathbb{R}$ be a general continuous taper function on an open interval I containing $[0,1]$ such that $h(0) = h(1) = 0$. For $\ell = 1, 2, \ldots$, denote by $\omega_\ell = 2\pi\ell/n$ the Fourier frequencies. The tapered DFT of $\{x_j\}_{j=-\infty}^{\infty}$ with taper function h is defined by

$$d_{x,\ell} = \sum_{j=1}^{n} h\Big(\frac{j}{n}\Big) x_j\, \mathrm{e}^{\mathrm{i}j\omega_\ell} = \sum_{j=1}^{n} h_\ell\Big(\frac{j}{n}\Big) x_j.$$

where $h_\ell(t) = h(t)\mathrm{e}^{2\pi\mathrm{i}\ell t}$. Denote by $\{\Delta x_j\}_{j=-\infty}^{\infty}$ the first difference of the series $\{x_j\}$, where $\Delta x_j = x_j - x_{j-1}$. We define the tapered DFT of the first difference $\{\Delta x_j\}_{j=-\infty}^{\infty}$ with taper function h by

$$d_{\Delta x,\ell} = \sum_{j=1}^{n} h\Big(\frac{j}{n}\Big) \Delta x_j\, \mathrm{e}^{\mathrm{i}j\omega_\ell} = \sum_{j=1}^{n} h_\ell\Big(\frac{j}{n}\Big) \Delta x_j. \tag{3.6}$$

In our setting, we observe the cointegrated component processes y_1 and y_2 at equidistant sample points. Defining the cointegrating error $z_j = y_{1,j} - \theta y_{2,j}$ and following Chen and Hurvich (2003b), we can now introduce the estimator

$$\hat{\theta}_n^{Tap} = \mathrm{Re}(\tilde{\theta}_n)\ ,$$

where $\mathrm{Re}(z)$ signifies the real part of a complex number $z = a + \mathrm{i}b$ and, letting $\bar{z} = a - \mathrm{i}b$ be the complex conjugate of z,

$$\tilde{\theta}_n = \frac{\sum_{\ell=1}^{m} d_{\Delta y_1,\ell}\, \bar{d}_{\Delta y_2,\ell}}{\sum_{\ell=1}^{m} |d_{\Delta y_2,\ell}|^2}\ .$$

Therein, any tapered DFT of differenced sequences is defined according to (3.6). Note that $\hat{\theta}_n^{Tap}$ is the real part of the ratio of the averaged tapered cross-periodogram between the series y_1 and y_2 and the averaged tapered periodogram of the series y_2.

3.3.1 Discrete tapered estimator under weak fractional cointegration

Theorem 3.4. *Let Assumptions 2.1, 2.2, 2.3, 2.5 and 2.6 hold. Assume moreover that the counting processes are mutually independent and independent of the efficient shocks and (3.3) holds. Then*

$$n^{1/2-H}(\hat{\theta}_n^{Tap} - \theta) \xrightarrow{d} \sqrt{\frac{c_1^2\lambda_1^{2H} + \theta^2 c_2^2 \lambda_2^{2H}}{\theta^{-2}\lambda_1\sigma_{1,e}^2\lambda_1 + \lambda_2\sigma_{2,e}^2}} \frac{\sum_{\ell=1}^m \mathrm{Re}\left(\int_0^1 h_\ell(s)\,\mathrm{d}B(s) \int_0^1 h_\ell(t)\,\mathrm{d}B_H(t)\right)}{\sum_{\ell=1}^m \left|\int_0^1 h_\ell(s)\,\mathrm{d}B(s)\right|^2}$$

where B is a standard Brownian motion, B_H is a standard fractional Brownian motion and B and B_H are independent.

Since the assumptions of Theorem 3.4 are the same as in Theorem 3.2, Remark 3.1 also applies here.

3.3.2 Discrete tapered estimator under strong fractional and standard cointegration

Theorem 3.5. *Let Assumptions 2.1, 2.2, 2.3 2.5 and 2.7 hold. Assume morover that the counting processes are mutually independent and independent of the efficient shocks and (3.4) holds.*

- *If $1/2 < H < 1$, then*

$$n^{3/2-H}(\hat{\theta}_n^{Tap} - \theta) \xrightarrow{d} \sqrt{\frac{c_1^2 + \theta^2 c_2^2}{\theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2}} \frac{\sum_{\ell=1}^m \mathrm{Re}\left(\int_0^1 h_\ell(s)\,\mathrm{d}B(s) \int_0^1 h'_\ell(s)\,\mathrm{d}B_H(s)\right)}{\sum_{\ell=1}^m |\int_0^1 h_\ell(s)\,\mathrm{d}B(s)|^2}$$

  where B_H is a standard fractional Brownian motion independent of the standard Brownian motion B.

- *If $H = 1/2$, $n(\hat{\theta}_n^{Tap} - \theta) = O_P(1)$.*

The assumptions of this theorem are weaker than those of Theorem 3.3 on the OLS estimator. The microstructure shocks are not assumed to be independent of the counting processes and the efficient shocks are not assumed to be Gaussian. Theorem 3.3 can presumably be proved without the Gaussian assumption. It might be much more difficult in the proof of Theorem 3.3 to avoid the assumption of independence between the microstructure shocks and the counting processes.

Remark 3.2 remains relevant here. In particular, we are currently unable to show that Assumption 2.7 holds except in certain cases of integer-valued durations. This motivates our consideration of a continuous-time tapered estimator in the following section.

3.4 A Continuous-Time Tapered Estimator

The estimators of θ we have considered so far are based on equally-spaced observations of the log price series. However, under the model (2.1), (2.2), a continuous-time record is available, and it is of interest to consider using all of the available data to estimate θ. Here, for the sake of theoretical tractability, we consider a tapered estimator $\tilde{\theta}$ based on continuously-averaged log prices on adjacent non-overlapping time intervals. Since the problems with discretization appear only in the strong fractional and standard cointegration cases, we only consider them in this section. There is no difference in the case of weak fractional cointegration.

We first establish some notation. Let $\{X(t)\}$ be any time series defined for all $t \geq 0$, and suppose that we have data on $\{X_t\}$ for $t \in [0, T]$. Let $\delta > 0$ be fixed. In practice, we might take δ to be 5 minutes, but the choice of δ does not affect the asymptotic distribution we derive below. Define $n = \lfloor T/\delta \rfloor$, $\tilde{X}(0) = 0$, and

$$\tilde{X}(k) = \int_{u=(k-1)\delta}^{k\delta} X(u)\,\mathrm{d}u\ ,\ \ k = 1, \cdots, n\ .$$

Then we can define an estimator $\tilde{\theta}_\delta$ based on these averaged observations by

$$\hat{\theta}_{n,\delta}^{Tap} = \mathrm{Re}(\tilde{\theta}_{n,\delta})$$

with

$$\tilde{\theta}_{n,\delta} = \frac{\sum_{\ell=1}^m d_{\Delta\tilde{y}_1,\ell}\bar{d}_{\Delta\tilde{y}_2,\ell}}{\sum_{\ell=1}^m |d_{\Delta\tilde{y}_2,\ell}|^2}\ .$$

3.4.1 Continuous-time tapered estimator under strong fractional and standard cointegration

Theorem 3.6. *Let Assumptions 2.1, 2.2, 2.3, 2.5 and 2.8 hold. Assume morover that the counting processes are mutually independent and independent of the efficient shocks and (3.4) holds.*

- *If $1/2 < H < 1$, then*

$$n^{3/2-H}(\hat{\theta}_{n,\delta}^{Tap} - \theta) \xrightarrow{d} \sqrt{\frac{\delta^{2H}(c_1^2 + \theta^2 c_2^2)}{\theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2}}\frac{\sum_{\ell=1}^m \mathrm{Re}\left(\int_0^1 h_\ell(s)\,\mathrm{d}B(s)\ \int_0^1 h'_\ell(s)\mathrm{d}B_H(s)\right)}{\sum_{\ell=1}^m |\int_0^1 h_\ell(s)\,\mathrm{d}B(s)|^2}$$

  where B_H is a standard FBM independent of the standard Brownian motion B.

- *If $H = 1/2$, $n(\hat{\theta}_{n,\delta}^{Tap} - \theta) = O_P(1)$.*

Because Assumption 2.8 involves an integral rather than a sum, we are able to verify that it holds for certain models with noninteger durations such as ACD and LMSD under certain relationships with the microstructure shocks.

In Theorem 3.6, we allow for leverage effects, and therefore care is required in defining standard cointegration. As demonstrated in Lemma 2.2 (which assumes LMSD durations) if there is a leverage effect, even when the microstructure shocks are the differences of a weakly-dependent sequence, the cointegrating error need not be $I(0)$. In such a case we have strong fractional cointegration rather than the standard cointegration which might have been anticipated.

It is also possible that even though a leverage effect exists, the memory of durations has no effect on the degree of cointegration. Specifically, if in Lemma 2.2 we replace $\xi_k = Y_{k+1}^2 - 1$ by $\xi_k = H_2(Y_{k+1}) - .75H_3(Y_{k+1})$, where $H_2(y) = y^2 - 1$ and $H_3(y) = y^3 - 3y$ (the second and third Hermite polynomials, respectively), then there is a leverage effect with $\mathrm{corr}(\tau_{k+1}, \xi_k) = .082$. Nevertheless it follows from an argument similar to the proof of Lemma 2.2 that Assumption 2.8 holds in this example with $H = 1/2$, so that we have standard cointegration and Theorem 3.6 holds with $H = 1/2$.

Lemma 4.10 provides an example of standard cointegration allowing for both time deformation and dependence between the counting processes and microstructure shocks. Theorem 3.6 would hold for this example with $H = 1/2$.

3.5 Spurious Regressions

In this subsection only, we consider a non-cointegrated version of the model defined by (2.1) and (2.2),

$$y_1(t) = \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \theta_{21} \sum_{k=1}^{N_2(t_{1,N_1(t)})} e_{2,k}, \tag{3.7}$$

$$y_2(t) = \sum_{k=1}^{N_2(t)} (e_{2,k} + \eta_{2,k}) + \theta_{12} \sum_{k=1}^{N_1(t_{2,N_2(t)})} e_{1,k}, \tag{3.8}$$

where $\theta_{12} \neq \theta_{21}^{-1}$. We examine here the properties of the OLS estimator in the (spurious) regression of y_1 on y_2 in discrete time and then briefly discuss corresponding tests for the null hypothesis of cointegration. Corollary 3.1 below follows directly from the proof of Theorem 3.1.

Corollary 3.1. *If Assumptions 2.1, 2.2, 2.3 and 2.4 are satisfied and $y = (y_1, y_2)$ is given by (3.7) and (3.8) with $\theta_{12} \neq \theta_{21}^{-1}$, then as $n \to \infty$,*

$$\left(\frac{1}{\sqrt{n}} y(nu) \colon u \in [0,1]\right) \Rightarrow B_y = (B_y(u) \colon u \in [0,1]),$$

where B_y is a bivariate Brownian motion with 2×2 covariance matrix $\Sigma = (\Sigma_{i,j}\colon i,j=1,2)$ given by the entries

$$\Sigma_{1,1} = \lambda_1\sigma_{1,e}^2 + \theta_{21}^2\lambda_2\sigma_{2,e}^2\ ,\ \ \Sigma_{2,2} = \theta_{12}^2\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2\ ,$$
$$\Sigma_{1,2} = \theta_{12}\lambda_1\,\sigma_{1,e}^2 + \theta_{21}\lambda_2\,\sigma_{2,e}^2 = \Sigma_{2,1}.$$

Next, we consider the discretization of $y_1(t)$ and $y_2(t)$ given by (3.7) and (3.8) at integer time values,

$$y_{1,j} = \sum_{k=1}^{N_1(j)}(e_{1,k}+\eta_{1,k}) + \theta_{21}\sum_{k=1}^{N_2(t_{1,N_1(j)})} e_{2,k}, \tag{3.9}$$
$$y_{2,j} = \sum_{k=1}^{N_2(j)}(e_{2,k}+\eta_{2,k}) + \theta_{12}\sum_{k=1}^{N_1(t_{2,N_2(j)})} e_{1,k}. \tag{3.10}$$

Regressing $y_{1,1},\ldots,y_{1,n}$ on $y_{2,1},\ldots,y_{2,n}$ without intercept, we obtain the OLS estimator

$$\hat{\delta}_n = \frac{\sum_{j=1}^n y_{2,j}y_{1,j}}{\sum_{j=1}^n y_{2,j}^2}. \tag{3.11}$$

Corollary 3.2 below follows directly from Corollary 3.1 and the Continuous Mapping Theorem.

Corollary 3.2. *If Assumptions 2.1, 2.2, 2.3 and 2.4 are satisfied and $y=(y_1,y_2)$ is given by (3.7) and (3.8) with $\theta_{12}\neq\theta_{21}^{-1}$, then as $n\to\infty$,*

$$\hat{\delta}_n \xrightarrow{d} \frac{\int_0^1 B_{2,y}(u)B_{1,y}(u)\,du}{\int_0^1 B_{1,y}^2(u)\,du}\ ,$$

where $B_y=(B_{1,y},B_{2,y})$ is the bivariate Brownian motion given in Corollary 3.1.

Corollary 3.2 together with Corollary 3.1 can be used to motivate tests for the null hypothesis of no cointegration. We do not pursue the details here, but it seems clear that the null distribution for unit root tests based on the residuals $\{y_{1,j}-\hat{\delta}_n y_{2,j}\}_{j=1}^n$ can be derived from Corollaries 3.1 and 3.2, and that these null distributions will have form similar to the distributions listed, for example, in Hamilton (1994, Proposition 19.4).

4 Proofs

Proof of Lemma 2.1. Since $\tilde{N}$ is stationary and ergodic, there exists $\tilde{\lambda}\in(,\infty)$ such that $\tilde{N}(t)/t \xrightarrow{P} \tilde{\lambda}$ a st $\to\infty$, a. s. See Daley and Vere-Jones (2003, diplays (12.2.3) and (12.2.4)).

Now,

$$\frac{N(t)}{t} = \frac{\tilde{N}(f(t))}{t} = \frac{\tilde{N}(f(t))}{f(t)} \frac{f(t)}{t} \; .$$

Since by assumption $f(t) \to \infty$ (in probability if f is random), the assumption implies that the first term on the righthand side converges to $\tilde{\lambda}$. By assumption, the second term satisfies $t^{-1} f(t) \xrightarrow{P} \gamma$. Thus, $\frac{N(t)}{t} \xrightarrow{P} \lambda$, so that Assumption 2.2 holds with $\lambda = \tilde{\lambda}\gamma$. We note that $N(t_k^-) \leq k$, thus

$$\begin{aligned} 1 \leq \frac{N(t_k)}{k} &= 1 + \frac{N(t_k) - k}{k} \leq 1 + \frac{N(t_k) - N(t_k^-)}{k} \\ &= 1 + \frac{\tilde{N}(f(t_k)) - \tilde{N}(f(t_k^-))}{k} \leq 1 + \frac{\tilde{N}(f(t_k)) - \tilde{N}(f(t_k) - C)}{k} \end{aligned}$$

using the definition of N and the boundedness requirement on the jumps of f. Since the process $\tilde{N}$ is stationary, taking expectations, we have

$$\mathbb{E}[\tilde{N}(f(t_k)) - \tilde{N}(f(t_k) - C)] = \mathbb{E}[\tilde{N}((0, C])] < \infty \; .$$

(In the case where f is random and independent of $\tilde{N}$ this holds by taking conditional expectation first). Thus it follows that

$$\frac{N(t_k)}{k} = 1 + O_P(1/k) \; .$$

Now, this implies that $N(t_k)/t_k$ converges in probability to 1. Thus

$$\frac{t_k}{k} = \frac{t_k}{N(t_k)} \frac{N(t_k)}{k} \xrightarrow{P} \frac{1}{\lambda} \; .$$

□

Proof of Lemma 2.2. Denote H_τ by H to simplify the notation. Define $\xi_k = Y_{k+1}^2 - 1$. Then ξ_k is centered, has finite variance summable autocovariance function, since $\mathrm{cov}(\xi_0, \xi_k) = 2\mathrm{cov}^2(Y_0, Y_{k+1})$. Thus $\{\xi_k\}$ has a summable autocovariance function because $H \in (1/2, 3/4)$. By Arcones (1994, Theorem 4), this implies that $\{\xi_k\}$ is in the domain of attraction of the standard Brownian motion, i.e.

$$n^{-1/2} \sum_{k=1}^{[n\cdot]} \xi_k \Rightarrow c' B \; ,$$

with $c'^2 = \mathrm{var}(\xi_0) + 2\sum_{k=1}^{\infty} \mathrm{cov}(\xi_0, \xi_k)$. This proves (2.6).

Assume now that $\tau_k = \epsilon_k \mathrm{e}^{\sigma Y_k}$ (with $\sigma = 1$ in the statement of the Lemma). The properties of Hermite polynomials yield that $\mathbb{E}[\mathrm{e}^{\sigma Y_0} H_j(Y_0)] = \sigma^j \mathrm{e}^{\sigma^2/2}$ for all $j \geq 1$. Denote

now $\lambda^{-1} = \mathbb{E}[\tau_k] = \mathbb{E}[\mathrm{e}^{\sigma Y_k}] = \mathrm{e}^{\sigma^2/2}$, $m = \mathbb{E}[\xi_{k-1}\tau_k] = \mathbb{E}[(Y_k^2-1)\mathrm{e}^{Y_k}] = \sigma^2\mathrm{e}^{\sigma^2/2}$ and $G(y) = (y^2-1)\mathrm{e}^{\sigma y} - m$. We now prove that (2.7) holds with $\mu^* = \lambda m$. Write

$$
\begin{aligned}
&\int_0^T (\xi_{N(s)} - \lambda m)\,\mathrm{d}s \\
&\quad = \sum_{k=0}^{N(T)} \tau_{k+1}\xi_k - \lambda m T + (t_{N(T)+1} - T)\xi_{N(T)+1} \\
&\quad = \sum_{k=0}^{N(T)} (\epsilon_{k+1}-1)\xi_k \mathrm{e}^{\sigma Y_{k+1}} + \sum_{k=0}^{N(T)} G(Y_{k+1}) + m(N(T) - \lambda T) - (t_{N(T)+1} - T)\xi_{N(T)+1}\ .
\end{aligned}
\tag{4.1}
$$

By Lemma 4.8 and applying Hölder's inequality, it can be shown that $(t_{N(T)+1}-T)(\xi_{N(T)+1}-\rho) = O_P(1)$. Since the sequence $\{\epsilon_k\}$ is independent of the Gaussian process $\{Y_k\}$, the first term in the righthand side of (4.1) is in the domain of attraction of the standard Brownian motion, and the normalizing sequence is $\sqrt{n}$. Thus we must obtain the joint asymptotic behaviour of $\sum_{k=1}^{N(Tt)} G(Y_k)$ and $N(Tt) - \lambda Tt$.

The durations are in the domain of attraction of the fractional Brownian motion with Hurst index H, since

$$
\sum_{k=1}^{n} (\tau_k - \lambda^{-1}) = \sum_{k=1}^{n} (\epsilon_k - 1)\mathrm{e}^{\sigma Y_k} + \sum_{k=1}^{n} (\mathrm{e}^{\sigma Y_k} - \lambda^{-1})\ .
$$

The first term in the righthand side is $O_P(\sqrt{n})$ and the second sum, suitably normalized converges to the fractional Brownian motion with Hurst index H because the function $x \to \mathrm{e}^{\sigma x} - \lambda^{-1}$ has Hermite rank 1. See e.g. Arcones (1994). More precisely, let $c_1 = \mathbb{E}[Y_1\mathrm{e}^{\sigma Y_1}] = \sigma\mathrm{e}^{\sigma^2/2}$ and define $g(y) = \mathrm{e}^{\sigma y} - \lambda^{-1} - c_1 y$. The function g has Hermite rank 2, and since $H \in (1/2, 3/4)$, this implies that

$$
\mathrm{var}\left(\sum_{k=1}^{n} g(Y_k)\right) = O(n)\ .
$$

Thus $\sum_{k=1}^{n}(\tau_k - \lambda^{-1})$ is asymptotically equivalent to $c_1\sum_{k=1}^{n} Y_k$. Let B_H denote the standard fractional Brownian motion with hurst index H. The assumption on the covariance of the Gaussian process $\{y_k\}$ implies that

$$
n^{-H}\sum_{k=1}^{[n\cdot]} Y_k \Rightarrow \varphi B_H
$$

with $\varphi^2 = c/\{H(2H-1)\}$. Denote now $c_2 = \mathbb{E}[Y_1 G(Y_1)] = \sigma(\sigma^2+2)\mathrm{e}^{\sigma^2/2}$ and define $h(y) = G(y) - c_2 y$. Then h has Hermite rank 2 and thus by similar arguments as above,

$\sum_{k=1}^{n} G(Y_k)$ is asymptotically equivalent to $c_2 \sum_{k=1}^{n} Y_k$. Thus we obtain

$$n^{-H} \left(\sum_{k=1}^{[nt]} (\tau_k - \lambda^{-1}), \sum_{k=1}^{[nt]} G(Y_k) \right) \Rightarrow (c_1 \varphi B_H(t), c_2 \varphi B_H(t)) \ .$$

By Vervaat's Lemma (see Vervaat (1972) or Resnick (2007, Proposition 3.3)), the previous convergence implies that

$$n^{-H} \left(N(nt) - \lambda nt, \sum_{k=1}^{[nt]} G(Y_k) \right) \Rightarrow (-\lambda c_1 \varphi B_H(\lambda t), c_2 \varphi B_H(t)) \ .$$

By the continuity of the composition map, this yields

$$n^{-H} \left(N(nt) - \lambda nt, \sum_{k=1}^{N(nt)} G(Y_k) \right) \Rightarrow (-\lambda c_1 \varphi B_H(\lambda t), c_2 \varphi B_H(\lambda t)) \ .$$

Next we obtain that

$$n^{-H} \left\{ \sum_{k=1}^{N(nt)} G(Y_k) + m(N(nt) - \lambda nt) \right\} \Rightarrow \varphi (c_2 - \lambda m c_1) B_H(\lambda t)$$

with $c_2 - \lambda m c_1 = 2\sigma \mathrm{e}^{\sigma^2/2} > 0$. We conclude that $n^{-H} \int_0^{n\cdot} \{\xi_{N(s)} - \lambda m\} \, \mathrm{d}s \Rightarrow \varphi(c_2 - \lambda m c_1) B_H$. □

4.1 Proof of Theorem 3.1 and Corollary 3.1

We first need the following Lemma.

Lemma 4.1. *Under Assumption 2.1 and 2.2, $N_i(t_{j,N_j(nt)})/n$ converges in probability uniformly on compact sets to $\lambda_i t$, where $\{i, j\} = \{1, 2\}$.*

Proof of Lemma 4.1. The sequence of (random) functions $N_i(n\cdot)/n$ is nondecreasing and converges pointwise in probability to $\lambda_i t$ by ergodicity. A sequence of nondecreasing function converging to a continuous function converges uniformly on compact sets. This results is known as Dini's Theorem. Cf. Resnick (1987, page 3). Thus the convergence of $N_i(n\cdot)/n$ is uniform on compact sets. Assumptions 2.1 and 2.2 imply that $N_i(t) \xrightarrow{P} \infty$ and $t_{i,n} \xrightarrow{P} \infty$. Thus

$$\frac{N_i(t_{j,N_j(nu)})}{n} = \frac{N_i(t_{j,N_j(nu)})}{t_{j,N_j(nu)}} \times \frac{t_{j,N_j(nu)}}{N_j(nu)} \times \frac{N_j(nu)}{n} \xrightarrow{P} \lambda_i \times \frac{1}{\lambda_j} \times \lambda_j u = \lambda_i u \ .$$

Applying again Dini's lemma, we also have that $N_i(t_{j,N_j(nu)})/n$ converges uniformly on compact sets to $\lambda_i u$. □

Proof of Theorem 3.1. Denote $S^e_{i,n}(t) = \sum_{k=1}^{[nt]} e_{i,k}$ and $S^\eta_{i,n}(t) = \sum_{k=1}^{[nt]} \eta_{i,k}$, $i = 1, 2$. Under Assumptions 2.3 and 2.4, $n^{-1/2}(S^e_{1,n}, S^e_{2,n}, S^\eta_{1,n}, S^\eta_{1,n})$ converges weakly to $(\sigma_{1,e}B_1, \sigma_{2,e}B_2, 0, 0)$, where B_1 and B_2 are independent standard Brownian motions. This follows from the independence of e_1 and e_2 and the local uniform convergence to 0 in probability of $n^{-1/2}S^\eta_{i,n}$. With the previous notation, (3.7) and (3.8) become

$$\begin{aligned} y_1(nt) &= S^e_{1,n}(N_1(nt)) + \theta_{21}S^e_{2,n}(N_2(t_{1,N_1(nt)})) + S^\eta_{1,n}(N_1(nt)) \;, \\ y_2(nt) &= S^e_{2,n}(N_2(nt)) + \theta_{12}S^e_{1,n}(N_1(t_{2,N_2(nt)})) + S^\eta_{2,n}(N_2(nt)) \;. \end{aligned}$$

By Lemma 4.1 and the continuity of the composition map on $\mathcal{C} \times \mathcal{C}$ endowed with the metric of uniform convergence on compact sets (see e.g. Billingsley (1968, Chapter 3, Section 17)), we obtain the joint convergence of

$$\begin{aligned} n^{-1/2} \big(& S^e_{1,n}(N_1(n\cdot)), S^e_{1,n}(N_1(t_{2,N_2(n\cdot)})), \\ & S^e_{2,n}(N_2(n\cdot)), S^e_{2,n}(N_2(t_{1,N_1(n\cdot)})), S^\eta_{1,n}(N_1(n\cdot)), S^\eta_{2,n}(N_2(n\cdot)) \big) \end{aligned}$$

to $(\sigma_{1,e}\sqrt{\lambda_1}B_1, \sigma_{1,e}\sqrt{\lambda_2}B_1, \sigma_{2,e}\sqrt{\lambda_2}B_2, \sigma_{2,e}\sqrt{\lambda_2}B_2, 0, 0)$. This yields Corollary 3.1 and Theorem 3.1 by setting $\theta_{21} = \theta$ and $\theta_{12} = \theta^{-1}$. $\square$

4.2 Proof of Theorems 3.2 and 3.3

Proof of Theorem 3.2. Write

$$\tilde\theta_n = \theta + \frac{\sum_{j=1}^n \{y_1(j) - \theta y_2(j)\} y_2(j)}{\sum_{j=1}^n y_2^2(j)} \;.$$

Assumptions 2.1, 2.2, 2.3, 2.5 and 2.6 imply those of Theorem 3.1. Thus we can apply the Continuous Mapping Theorem and obtain

$$n^{-2} \sum_{j=1}^n y_2^2(j) \xrightarrow{d} \{\theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2\} \int_0^1 B^2(s)\,\mathrm{d}s \;, \tag{4.2}$$

where B is a standard Brownian motion. Thus, in order to study the convergence of $\tilde\theta_n - \theta$ suitably renormalized, it suffices to study the sum

$$\sum_{j=1}^n \{y_1(j) - \theta y_2(j)\} y_2(j) \;.$$

We further decompose the cointegrating error. Denote

$$
\begin{aligned}
y_1^e(j) &= \sum_{k=1}^{N_1(j)} e_{1,k} + \theta \sum_{k=1}^{N_2(t_{1,N_1(j)})} e_{2,k} \ , \quad y_1^\eta(j) = \sum_{k=1}^{N_1(j)} \eta_{1,k} \ , \\
y_2^e(j) &= \sum_{k=1}^{N_2(j)} e_{2,k} + \theta^{-1} \sum_{k=1}^{N_1(t_{2,N_2(j)})} e_{1,k} \ , \quad y_2^\eta(j) = \sum_{k=1}^{N_2(j)} \eta_{2,k} \ , \\
r_{1,j} &= \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} \ , \quad r_{2,j} = \sum_{k=N_2(t_{1,N_1(j)})+1}^{N_2(j)} e_{2,k} \ .
\end{aligned}
$$

With this notation, we can write

$$
\sum_{j=1}^{n} \{y_1(j) - \theta y_2(j)\} y_2(j) = \sum_{j=1}^{n} \{r_{1,j} - \theta r_{2,j}\} y_2(j) + \sum_{j=1}^{n} \{y_1^\eta(j) - \theta y_2^\eta(j)\} y_2(j) \ . \tag{4.3}
$$

Applying Theorem 3.1, Assumption 2.6 and the Continuous Mapping Theorem, we obtain

$$
\begin{aligned}
& n^{-3/2-H} \sum_{j=1}^{n} \{y_1^\eta(j) - \theta y_2^\eta(j)\} y_2(j) \\
& \quad \xrightarrow{d} \int_0^1 \{\theta^{-1} \sqrt{\lambda_1} \sigma_{1,e} B_1(t) + \sqrt{\lambda_2} \sigma_{2,e} B_2(t)\} \{c_1 B_{1,H}(\lambda_1 t) - \theta c_2 B_{2,H}(\lambda_2 t)\} \, \mathrm{d}t \\
& \quad \stackrel{d}{=} \Sigma \int_0^1 B(t) B_H(t) \, \mathrm{d}t
\end{aligned}
$$

where B is a standard Brownian motion, B_H is a fractional Brownian motion, independent of B and

$$
\Sigma^2 = (\theta^{-2} \lambda_1 \sigma_{1,e}^2 + \lambda_2 \sigma_{2,e}^2)(c_1^2 \lambda_1^{2H} + \theta^2 c_2^2 \lambda_2^{2H}) \ . \tag{4.4}
$$

There only remains to prove that, for $i = 1, 2$,

$$
n^{-3/2} \sum_{j=1}^{n} r_{i,j} y_2(j) = O_P(1) \ . \tag{4.5}
$$

The convergence of $n^{-1/2} y_2$ is uniform on $[0, 1]$, so $n^{-1/2} \max_{1 \le j \le n} |y_2(j)| = O_P(1)$. Therefore, it suffices to prove that

$$
n^{-1} \sum_{j=1}^{n} |r_{i,j}| = O_P(1) \ . \tag{4.6}
$$

Recall that $N_i(s) < k \Leftrightarrow t_{i,k} > s$. Thus, for $k \le N_1(n)$,

$$
N_1(t_{2,N_2(j)}) < k \le N_1(j) \Leftrightarrow t_{2,N_2(j)} < t_{1,k} \le j \ .
$$

The first inequality on the righthand side means that there is no point of N_2 between $t_{1,k}$ and j, i.e. $j \le t_{2,N_2(t_{1,k})+1}$. Let $A_2(t) = t_{2,N_2(t)+1} - t$ denote the forward recurrence time of N_2, i.e. the time between t and the next event of N_2 after t. Thus,

$$\sum_{j=1}^n |r_{1,j}| \le \sum_{j=1}^n \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} |e_{1,k}| = \sum_{k=1}^{N_1(n)} |e_{1,k}|\{A_2(t_{1,k}) + 1\} \; .$$

We thus get the bound for the conditional expectation given the sigma-field $\mathcal{N}$ generated by the counting processes N_1 and N_2:

$$\mathbb{E}\left[\sum_{j=1}^n |r_{1,j}| \mid \mathcal{N}\right] \le C \sum_{k=1}^{N_1(n)} A_2(t_{1,k}) \; .$$

Conditioning on N_1 and applying (3.3) yields

$$\mathbb{E}\left[\sum_{j=1}^n |r_{1,j}| \mid N_1\right] \le C N_1(n) = O_P(n) \; .$$

This proves (4.6) and concludes the proof of Theorem 3.2. □

Proof of Theorem 3.3. The proof is a consequence of the convergence (4.2), the decomposition (4.3), and Lemmas 4.2 and 4.3, whose assumptions are those of the Theorem. □

Lemma 4.2. *Under the assumptions of Theorem 3.3,*

$$n^{-H-1/2} \sum_{j=1}^n \{y_1^\eta(j) - \theta y_2^\eta(j)\} y_2(j) \xrightarrow{d} \Sigma_0 \int_0^1 B(s) \, \mathrm{d}B_H(s) \; . \tag{4.7}$$

where B_H is a standard fractional Brownian motion independent of B and

$$\Sigma_0 = (\theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2)(c_1^2 + \theta^2 c_2^2) \; .$$

Proof of Lemma 4.2. Denote $S_n = \sum_{j=1}^n \{y_1^\eta(j) - \theta y_2^\eta(j)\} y_2(j)$ and write $y_2 = y_2^e + y_2^\eta$ with obvious notation. Denote $\zeta_j = y_1^\eta(j) - \theta y_2^\eta(j) = \xi_{1,N_1(j)} - \theta \xi_{2,N_2(j)}$. Then

$$S_n = \sum_{j=1}^n \zeta_j y_2^e(j) + \sum_{j=1}^n \zeta_j \xi_{2,N_2(j)} \; . \tag{4.8}$$

By the last part of Assumption 2.7, the last term in the righthand side of (4.8) is $O_P(n)$. Consider the first term in the righthand side of (4.8), say $S_{1,n}$. Write

$$\begin{aligned} S_{1,n} &= \sum_{j=1}^n \zeta_j \sum_{k=1}^{N_2(j)} e_{2,k} + \theta^{-1} \sum_{j=1}^n \zeta_j \sum_{k=1}^{N_1(t_{2,N_2(j)})} e_{1,k} \\ &= \sum_{k=1}^{N_2(n)} e_{2,k} \sum_{\{j \le n :\, N_2(j) \ge k\}} \zeta_j + \theta^{-1} \sum_{k=1}^{N_1(t_{2,N_2(n)})} e_{1,k} \sum_{\{j \le n :\, N_1(t_{2,N_2(j)}) \ge k\}} \zeta_j \\ &= T_{1,n} + \theta^{-1} T_{2,n} \; . \end{aligned}$$

Denote $W_n(t) = \sum_{j=1}^{[nt]} \zeta_j$. Since $N_2(j) < k$ iff $j < t_{2,k}$, we obtain

$$T_{1,n} = y_2^{e_2}(n) W_n(1) - \sum_{k=1}^{N_2(n)} e_{2,k} W_n(t_{2,k}/n) \ .$$

By Assumption 2.7 and Theorem 3.1, $n^{-1/2-H} y_2^{e_2}(n) W_n(1) \xrightarrow{d} \sqrt{\lambda_2}\sigma_2 B_2(1) Z(1)$ with $Z = c_1 B_H^{(1)} - \theta c_2 B_H^{(2)} \stackrel{d}{=} \sqrt{c_1^2 + \theta^2 c_2^2}\, B_H$. Let the last term be denoted by U_n. Since the shocks $e_{i,k}$ are i.i.d. Gaussian, we can compute the characteristic function of U_n.

$$\begin{aligned}\mathbb{E}[\exp\{itn^{-1/2-H} U_n\}] &= \mathbb{E}\left[\exp\left\{-\frac{\sigma_{2,e}^2 t^2}{2} \frac{1}{n} \sum_{k=1}^{N_2(n)} \left(n^{-H} W_n(t_{2,k}/n)\right)^2\right\}\right] \\ &\to \mathbb{E}\left[\exp\left\{-\frac{\lambda_2 \sigma_{2,e}^2 t^2}{2} \int_0^1 Z^2(s)\, \mathrm{d}s\right\}\right] \ .\end{aligned}$$

The convergence is actually joint with that of $n^{-1/2-H} y_2^{e_2}(n) W_n$, thus we have

$$n^{-1/2-H} T_{1,n} \xrightarrow{d} \sqrt{\lambda_2}\sigma_{2,e} B_2(1) Z(1) - \sqrt{\lambda_2}\sigma_{2,e} \int_0^1 Z(s)\, \mathrm{d}B_2(s) \ .$$

The limit can also be written as $\sqrt{\lambda_2}\sigma_{2,e} \int_0^1 B_2(s)\, \mathrm{d}Z(s)$. Consider now the term $T_{2,n}$. Note that $N_1(t_{2,N_2(j)}) < k$ iff $j \le t_{2,N_2(t_{1,k})+1}$. Thus

$$T_{2,n} = \sum_{k=1}^{N_1(t_{2,N_2(n)})} e_{1,k} W_n(1) - \sum_{k=1}^{N_1(t_{2,N_2(n)})} e_{1,k} W_n(t_{2,N_2(t_{1,k})+1}/n) \ .$$

By similar arguments as previously, we obtain

$$n^{-H-1/2} T_{2,n} \xrightarrow{d} \sqrt{\lambda_1}\sigma_1 B_1(1) Z(1) - \sqrt{\lambda_1}\sigma_1 \int_0^1 Z(s)\, \mathrm{d}B_1(s) \ .$$

All convergences hold jointly, thus (4.7) holds. $\square$

Lemma 4.3. *Under the assumptions of Theorem 3.3,*

$$\sum_{j=1}^{n} \{r_{1,j} - \theta r_{2,j}\} y_2(j) = O_P(n) \ , \tag{4.9}$$

Proof of Lemma 4.3. We first study the term with $r_{1,j}$ and split it into three parts.

$$\sum_{j=1}^{n} r_{1,j} y_2(j) = \sum_{j=1}^{n} r_{1,j} y_2^{e_1}(j) + \sum_{j=1}^{n} r_{1,j} y_2^{e_2}(j) + \sum_{j=1}^{n} r_{1,j} y_2^{\eta}(j)$$

We start with the last one. Recall that $N_1(t_{2,N_2(j)}) < k \le N_1(j)$ iff $t_{1,k} \le j \le t_{1,k} + A_2(t_{1,k})$. Thus

$$\sum_{j=1}^{n} r_{1,j} y_2^{\eta}(j) = \sum_{j=1}^{n} \xi_{2,N_2(j)} \sum_{N_1(t_{2,N_2(j)})<k\le N_1(j)} e_{1,k} \ . \tag{4.10}$$

If the microstructure shocks are independent of the counting processes, then

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{j=1}^{n} e_{1,k} \sum_{t_{1,k}\le j<t_{1,k}+A_2(t_{1,k})} \xi_{2,N_2(j)}\right)^2 \mid \mathcal{N}\right] \\
&= \sigma_{1,e}^2 \sum_{k=1}^{N_1(n)} \mathbb{E}\left[\left(\sum_{t_{1,k}\le j<t_{1,k}+A_2(t_{1,k})} \xi_{2,N_2(j)}\right)^2 \mid \mathcal{N}\right] \le C \sum_{k=1}^{N_1(n)} (A_2(t_{1,k})+1)^2 \sup_{\ell} \mathbb{E}[\xi_{2,\ell}^2] \ .
\end{aligned}$$

Conditioning on N_1 and then taking expectation yields

$$\mathbb{E}\Big[\Big(\sum_{j=1}^{n} e_{1,k} \sum_{t_{1,k}\le j<t_{1,k}+A_2(t_{1,k})} \xi_{2,N_2(j)}\Big)^2\Big] \le C\mathbb{E}[N_1(n)] \sup_t \mathbb{E}[\{1+A_2(t)\}^2] \sup_{\ell} \mathbb{E}[\xi_{\ell}^2] = O(n) \ .$$

Consider now $R_{2,n} = \sum_{j=1}^{n} r_{1,j} y_2^{e_2}(j)$.

$$R_{2,n} = \sum_{j=1}^{n} y_2^{e_2}(j) \sum_{N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} = \sum_{k=1}^{N_1(n)} e_{1,k} \sum_{t_{1,k}\le j<t_{1,k}+A_2(t_{1,k})} y_2^{e_2}(j) \ .$$

By independence of the efficient shocks and the counting processes, we have

$$\mathbb{E}[R_{2,n}^2 \mid \mathcal{N}] \le CN_1(n) \sum_{k=1}^{N_1(n)} (A_2(t_{1,k})+1)^2 = O_P(n^2) \ .$$

This proves that $R_{2,n} = O_P(n)$. Consider finally $R_{1,n} = \sum_{j=1}^{n} r_{1,j} y_2^{e_1}(j)$. By definition, $e_{1,k}$ is independent of $y_2^{e_1}(j)$ for j such that $N_1(t_{2,N_2(j)}) < k$. Thus, we can compute the conditional variance given $\mathcal{N}$.

$$\begin{aligned}
\mathbb{E}[R_{1,n}^2 \mid \mathcal{N}] &= \sigma_{1,e}^2 \sum_{k=1}^{N_1(n)} \mathbb{E}\left[\left(\sum_{t_{1,k}\le j<t_{1,k}+A_2(t_{1,k})} y_2^{e_1}(j)\right)^2 \mid \mathcal{N}\right] \\
&\le CN_2(n) \sum_{k=1}^{N_1(n)} (A_2(t_{1,k})+1)^2 = O_P(n)
\end{aligned}$$

by (3.4). This concludes the proof of Lemma 4.3. $\square$

4.3 Proof of Theorems 3.4 and 3.5

Write

$$\tilde{\theta}_n = \theta + \frac{\sum_{\ell=1}^m d_{\Delta r,\ell}\,\bar{d}_{\Delta y_2,\ell}}{\sum_{\ell=1}^m |d_{\Delta y_2,\ell}|^2} + \frac{\sum_{\ell=1}^m d_{\Delta y^\eta,\ell}\,\bar{d}_{\Delta y_2,\ell}}{\sum_{\ell=1}^m |d_{\Delta y_2,\ell}|^2}$$

with $y^\eta(j) = y_1^\eta(j) - \theta y_2^\eta(j)$, $r(j) = r_1(j) - \theta r_2(j)$ and

$$r_1(j) = \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k}\ , \quad r_2(j) = \sum_{k=N_2(t_{1,N_1(j)})+1}^{N_2(j)} e_{2,k}\ .$$

By summation by parts, since $h(0) = h(1) = 0$, for any time series $\{x_j\}$, we can write

$$d_{\Delta x,\ell} = \sum_{j=0}^{n-1} \{h_\ell(j/n) - h_\ell((j+1)/n)\} x_j = -\frac{1}{n}\sum_{j=0}^{n-1} w_\ell(j,n) x_j \tag{4.11}$$

with $w_\ell(j,n) = n\{h_\ell((j+1)/n) - h_\ell(j/n)\}$. Applying (4.11) to y_2 yields

$$d_{\Delta y_2,\ell} = -\frac{1}{n}\sum_{j=0}^{n-1} w_\ell(j,n) y_2(j)\ .$$

Since the assumptions of Theorems 3.4 and 3.5 imply those of Theorem 3.1, the Continuous Mapping Theorem yields

$$\{n^{-1/2} d_{\Delta y_2,\ell}, 1 \le \ell \le m\} \xrightarrow{d} \left\{-\Sigma_e \int_0^1 h_\ell'(s) B(s)\,\mathrm{d}s\ , 1 \le \ell \le m\right\} \tag{4.12}$$

where B is a standard Brownian motion and $\Sigma_e^2 = \theta^{-2}\lambda_1\sigma_{1,e}^2 + \lambda_2\sigma_{2,e}^2$. By integration by parts, the integral can also be expressed as

$$-\int_0^1 h_\ell'(s) B(s)\,\mathrm{d}s = \int_0^1 h_\ell(s)\,\mathrm{d}B(s)\ .$$

This in turn implies

$$n^{-1}\sum_{\ell=1}^m |d_{\Delta y_2,\ell}|^2 \xrightarrow{d} \Sigma_e^2 \sum_{\ell=1}^m \left|\int_0^1 h_\ell(s)\,\mathrm{d}B(s)\right|^2\ . \tag{4.13}$$

Applying now (4.11) to y^η we obtain

$$d_{\Delta y^\eta,\ell} = -\frac{1}{n}\sum_{j=0}^{n-1} w_\ell(j,n)\{y_1^\eta(j) - \theta y_2^\eta(j)\}\ .$$

In the case of weak fractional cointegration, we apply Assumption 2.6, the Continuous Mapping Theorem and integration by parts to obtain

$$n^{-H}\ell(n)d_{\Delta y^\eta,\ell} = -n^{-1-H}\ell(n)\sum_{j=0}^{n-1} w_\ell(j,n)\{y_1^\eta(j) - \theta y_2^\eta(j)\} \to \int_0^1 h_\ell(t)\,\mathrm{d}Z_H(t) \tag{4.14}$$

where, by independence of $B_H^{(1)}$ and $B_H^{(2)}$,

$$Z_H(t) = c_1 B_H^{(1)}(\lambda_1 t) - \theta c_2 B_H^{(2)}(\lambda_2 t) \overset{(law)}{=} \sqrt{\lambda_1^{2H}c_1^2 + \lambda_2^{2H}\theta^2 c_2^2}\, B_H$$

and B_H is a standard fractional Brownian motion. The first part of Lemma 4.4 shows that $d_{\Delta r,\ell}$ is negligible under the assumptions of Theorem 3.4. This, and the convergences (4.12), (4.13) and (4.14) conclude the proof of Theorem 3.4. $\square$

We now prove Theorem 3.5. Since $h_\ell(0) = h_\ell(1) = 0$, we have $\sum_{j=0}^{n-1} w_\ell(j,n) = 0$, hence

$$\sum_{j=0}^{n-1} w_\ell(j,n) y_i^\eta(j) = \sum_{j=0}^{n-1} w_\ell(j,n)(\xi_{i,N_i(j)} - \xi_{i,0}) = \sum_{j=0}^{n-1} w_\ell(j,n)(\xi_{i,N_i(j)} - \mu_i^*)\ .$$

Denote $S_{i,0} = 0$ and for $k \geq 1$, $S_{i,k} = \sum_{j=1}^{k}(\xi_{i,N_i(j)} - \mu_i^*)$. Define $\omega_\ell(j,n) = n\{w_\ell(j+1,n) - w_\ell(j,n)\}$. Applying again summation by parts, we have

$$\sum_{j=0}^{n-1} w_\ell(j,n) y_i^\eta(j) = -\frac{1}{n}\sum_{j=1}^{n-1}\omega_\ell(j,n)S_{i,j} + w_\ell(n,n)S_{i,n-1} + w_\ell(0,n)(\xi_{i,0} - \mu_i^*)\ ,$$

Under Assumption 2.7, by the Continuous Mapping Theorem, we obtain

$$\begin{aligned} n^{1-\gamma}\ell(n)d_{\Delta y^\eta,\ell} &= -n^{-\gamma}\ell(n)\sum_{j=1}^{n} w_\ell(j,n)y_i^\eta(j) \\ &\overset{d}{\to} \int_0^1 h_\ell''(t)B_H^{(i)}(t)\,\mathrm{d}t - h'(1)B_H^{(i)}(1) \overset{d}{=} -\int_0^1 h_\ell'(s)\,\mathrm{d}B_H^{(i)}(s)\ . \end{aligned} \tag{4.15}$$

The second part of Lemma 4.4 implies that the term $d_{\Delta r,\ell}$ does not contribute to the limit under the Assumptions of Theorem 3.5. This, and the convergences (4.12), (4.13) and (4.15) conclude the proof of Theorem 3.5. $\square$

Lemma 4.4. *Under the assumptions of Theorem 3.4, then $d_{\Delta r,\ell} = O_P(1)$. Under the assumptions of Theorem 3.5, then $d_{\Delta r,\ell} = O_P(n^{-1/2})$.*

Proof. Applying (4.11) to r, we see that we only need to prove that the independence between the counting processes and the efficient shocks and (3.3) implies that $\sum_{j=1}^{n} w_\ell(j,n)r_{i,j} = O_p(n)$ and (3.4) implies that $\sum_{j=1}^{n} w_\ell(j,n)r_{i,j} = O_p(n^{1/2})$. We start with r_1.

$$\sum_{j=1}^{n} w_\ell(j,n)r_{1,j} = \sum_{k=1}^{N_1(n)} e_{1,k} \sum_{t_{1,k}\leq j < t_{1,k}+A_2(t_{1,k})} w_\ell(j,n)\ .$$

Taking conditional expectation yields, for $q = 1, 2$,

$$\mathbb{E}\left[\left|\sum_{j=1}^{n} w_\ell(j,n) r_{1,j}\right|^q \mid \mathcal{N}\right] \le C \sum_{k=1}^{N_1(n)} (A_2(t_{1,k}) + 1)^q .$$

Applying (3.3) if $q = 1$ and (3.4) if $q = 2$ shows that the last term is $O_P(n)$. This proves that $\sum_{j=1}^{n} w_\ell(j,n) r_{1,j} = O_P(n)$ under the assumptions of Theorem 3.4 and $O_P(\sqrt{n})$ under the assumptions of Theorem 3.5. The term $\sum_{j=1}^{n} w_\ell(j,n) r_{2,j}$ is dealt with similarly. □

4.4 Proof of Theorem 3.6

Write

$$\tilde{\theta}_{n,\delta} = \theta + \frac{\sum_{\ell=1}^{m} d_{\Delta\tilde{r},\ell} \bar{d}_{\Delta\tilde{y}_2,\ell}}{\sum_{\ell=1}^{m} |d_{\Delta\tilde{y}_2,\ell}|^2} + \frac{\sum_{\ell=1}^{m} d_{\Delta\tilde{y}^\eta,\ell} \bar{d}_{\Delta\tilde{y}_2,\ell}}{\sum_{\ell=1}^{m} |d_{\Delta\tilde{y}_2,\ell}|^2}$$

with $\tilde{y}(j) = \tilde{y}_1^\eta(j) - \theta \tilde{y}_2^\eta(j)$, $\tilde{r}(j) = \tilde{r}_1(j) - \theta \tilde{r}_2(j)$ and

$$r_1(s) = \sum_{k=N_1(t_{2,N_2(s)})+1}^{N_1(s)} e_{1,k} \; , \quad r_2(s) = \sum_{k=N_2(t_{1,N_1(s)})+1}^{N_2(s)} e_{2,k} \; .$$

and the DFT is defined as in (3.6). Applying summation by parts as in (4.11), we obtain

$$d_{\Delta\tilde{y}_2,\ell} = -\frac{1}{n} \sum_{j=0}^{n-1} w_\ell(j,n) \tilde{y}_2(j) = -\frac{1}{n} \int_0^{n\delta} w_\ell(\lceil s/\delta \rceil, n) y_2(s) \, \mathrm{d}s = -\int_0^{\delta} w_\ell(\lceil nt/\delta \rceil, n) y_2(ns) \, \mathrm{d}t \; ,$$

with $w_\ell(j,n) = n\{h_\ell((j+1)/n) - h_\ell(j/n)\}$ as before, and $\lceil t \rceil$ is the smallest integer larger than or equal to t. This yields

$$n^{-1/2} d_{\tilde{y}_2,\ell} \xrightarrow{d} -\Sigma_e \int_0^1 h'_\ell(s) B(\delta s) \, \mathrm{d}s \stackrel{d}{=} \Sigma_e \int_0^1 h_\ell(s) \, \mathrm{d}B(s) \; .$$

Since $\eta_j = \xi_j - \xi_{j-1}$, we have

$$\tilde{y}_i^\eta(j) = \int_{(j-1)\delta}^{j\delta} \xi_{i,N_i(s)} \, \mathrm{d}s - \delta \xi_{i,0} \; .$$

Differencing cancels the term $\delta\xi_0$. Applying (4.11) and summation by parts and the property that $\sum_{j=0}^{n-1} w_\ell(j,n) = 0$, we obtain

$$\begin{aligned} d_{\Delta\tilde{y}_i^\eta,\ell} &= -\frac{1}{n} \sum_{j=0}^{n-1} w_\ell(j,n) \int_{(j-1)\delta}^{j\delta} \xi_{i,N_i(s)} \, \mathrm{d}s = -\frac{1}{n} \sum_{j=0}^{n-1} w_\ell(j,n) \int_{(j-1)\delta}^{j\delta} \{\xi_{i,N_i(s)} - \mu_i^*\} \, \mathrm{d}s \\ &= \frac{1}{n^2} \sum_{j=1}^{n-1} \omega_\ell(j,n) \int_0^{j\delta} \{\xi_{i,N_i(s)} - \mu_i^*\} \, \mathrm{d}s - \frac{1}{n} w_\ell(n,n) \int_0^{(n-1)\delta} \{\xi_{i,N_i(s)} - \mu_i^*\} \, \mathrm{d}s \; . \end{aligned}$$

Under Assumption 2.8, we thus have, with $Z = B_H^{(1)} - \theta B_H^{(2)}$,

$$n^{1-H}\{d_{\Delta\tilde{y}_1^\eta,\ell} - \theta d_{\Delta\tilde{y}_2^\eta,\ell}\} \xrightarrow{d} \int_0^1 h_\ell''(s) Z(\delta s)\,\mathrm{d}s - h'(1)Z(\delta)\ .$$

We must now deal with the remaining terms of the cointegrating error. If $H > 1/2$, Lemma 4.5 implies that the term $d_{\Delta\tilde{r},\ell}$ does not contribute to the limit. If $H = 1/2$, both terms are of the same order. This concludes the proof of Theorem 2.8. □

Lemma 4.5. *Under the assumptions of Theorem 3.6*

$$d_{\Delta\tilde{r}_i,\ell} = O_P(n^{-1/2})\ .$$

Proof. Applying as usual summation by parts, we obtain

$$\begin{aligned} d_{\Delta\tilde{r}_1,\ell} &= -\frac{1}{n}\sum_{k=1}^{N_1(n\delta)} e_{1,k}\sum_{j=1}^{n} w_\ell(j,n)\int_{(j-1)\delta}^{j\delta} \mathbf{1}_{\{t_{1,k}\le s<t_{1,k}+A_2(t_{1,k})\}}\,\mathrm{d}s \\ &= -\frac{1}{n}\sum_{k=1}^{N_1(n\delta)} e_{1,k}\int_0^{n\delta} w_\ell(\lceil s/\delta\rceil,n)\mathbf{1}_{\{t_{1,k}\le s<t_{1,k}+A_2(t_{1,k})\}}\,\mathrm{d}s \\ &= -\frac{1}{n}\sum_{k=1}^{N_1(n\delta)} e_{1,k}\int_{t_{1,k}\wedge(n\delta)}^{\{t_{1,k}+A_2(t_{1,k})\}\wedge(n\delta)} w_\ell(\lceil s/\delta\rceil,n)\,\mathrm{d}s \\ &= -\sum_{k=1}^{N_1(n\delta)} e_{1,k}\int_{(t_{1,k}/n)\wedge\delta}^{\{(t_{1,k}+A_2(t_{1,k}))/n\}\wedge\delta} w_\ell(\lceil nt/\delta\rceil,n)\,\mathrm{d}t\ . \end{aligned} \tag{4.16}$$

Taking conditional expectation and applying (3.4), we obtain

$$\mathbb{E}\left[|d_{\Delta\tilde{r}_1,\ell}|^2 \mid \mathcal{N}\right] \le \frac{C}{n^2}\sum_{k=1}^{N(n)} A_2^2(t_{1,k}) = O_P(n^{-1})\ .$$

□

4.5 Additional Lemmas

Lemma 4.6. *If the durations $t_{i,k} - t_{i,k-1}$ form a stationary ergodic sequence with finite moment of order $2p+1$, if $\mathbb{P}(t_{i,1} > 0) = 1$ and if the associated point process has finite intensity, then*

$$\sup_{s\ge 0}\mathbb{E}[(t_{i,N_i(s)+1} - s)^p] < \infty\ .$$

Proof of Lemma 4.6. We omit the index i. Let θ_t denote the shift operator and let $A(t)$ be the forward recurrence time. Then $A(s) = t_{N(s)+1} - s = t_1 \circ \theta_s$. Since the sequence $\{\tau_i\}$ is stationary under $\mathbb{P}$, there exists a probability law P^* such that N is a stationary ergodic point process under P^*, see Baccelli and Brémaud (2003, Section 1.3.5). Applying Baccelli and Brémaud (2003, Formula 1.3.3), we obtain

$$\begin{aligned}
\mathbb{E}[A^p(s)] &= \lambda^{-1}\mathbb{E}^*\left[\sum_{k=1}^{N(1)} t_1^p \circ \theta_s \circ \theta_{t_k}\right] = \lambda^{-1}\mathbb{E}^*\left[\sum_{k=1}^{N(1)} A^p(s+t_k)\right] \\
&= \lambda^{-1}\mathbb{E}^*\left[\sum_{k=1}^{N(1)} \{t_{N(s+t_k)+1} - s - t_k\}^p\right] \le \lambda^{-1}\mathbb{E}^*\left[\sum_{k=1}^{N(1)} \{t_{N(s+1)+1} - s\}^p\right] \\
&= \lambda^{-1}\mathbb{E}^*[N(1)\{t_{N(s+1)+1} - s\}^p] \le \lambda^{-1}\{\mathbb{E}^*[N(1)^2]\}^{1/2}\{\mathbb{E}^*[(t_{N(s+1)+1} - s)^{2p}]\}^{1/2} .
\end{aligned} \tag{4.17}$$

Since N is stationary under P^*, the last term does not depend on s, and by the Ryll-Nardzewski inversion formula (Baccelli and Brémaud (2003, Formula 1.2.25)), we have

$$\mathbb{E}^*[(t_{N(s+1)+1} - s)^{2p}] = \mathbb{E}^*[(t_1+1)^{2p}] = \lambda\mathbb{E}[\int_0^{t_1} (t_1 + 1 - s)^{2p}\, \mathrm{d}s \le \lambda\mathbb{E}[(1+t_1)^{2p+1}]$$

By Baccelli and Brémaud (2003, Property 1.6.3), the point process N is stationary and ergodic under P^* since the sequence of durations τ_k is stationary and ergodic. Thus, By Daley and Vere-Jones (2003, Theorem 3.5.III), $\mathbb{E}^*[N(0,1)^2] < \infty$. Plugging the last two bounds into (4.17), we obtain that $\mathbb{E}[A^p(s)]$ is uniformly bounded. □

Lemma 4.7. *Assume that there exists an increasing sequence $\{s_n, n \ge 0\}$ such that $s_0 = 0$ and*

(a) f is either constant or strictly increasing and differentiable on (s_n, s_{n+1}) and the jumps of f occur at some (but not necessarily all) of the s_n;

(b) if f is eiter constant or increasing on both intervals (s_n, s_{n+1}) and (s_{n+1}, s_{n+2}), then f has a jump at s_{n+1}.

Assume moreover that

- (minimum duration of trading and nontrading periods) *there exists $\delta_0 > 0$ such that $s_{n+1} - s_n \ge \delta_0$ for all $n \ge 0$;*
- (maximum duration of nontrading periods) *there exists C_0 such that for all $n \ge 0$, if f is constant on (s_n, s_{n+1}), then $s_{n+1} - s_n \le C_0$;*
- (non stoppage of time during trading periods) *there exists $\delta_1 > 0$ such that for all $n \ge 0$, f is either constant on (s_n, s_{n+1}), or $f'(t) \ge \delta_1$ for all $t \in (s_n, s_{n+1})$.*

Let $\tilde{N}$ be a point process with event times $\{\tilde{t}_k\}$ and let N be the point process defined by $N(\cdot) = \tilde{N}(f(\cdot))$ with event times $\{t_k\}$. If $\sup_{s\geq 0} \mathbb{E}[(\tilde{t}_{\tilde{N}(s)+1} - s)^p] < \infty$, then $\sup_{s\geq 0} \mathbb{E}[(t_{N(s)+1} - s)^p] < \infty$.

Proof of Lemma 4.7. Define the nondecreasing left-continuous inverse $f^{\leftarrow}$ of a nondecreasing càdlàg function f by

$$f^{\leftarrow}(u) = \inf\{t \mid f(t) \geq u\} .$$

Note first that $f^{\leftarrow}(u) \leq t$ if only if $u \leq f(t)$ and $f^{\leftarrow}(f(t)) \leq t$. Thus we see that

$$\begin{aligned} f^{\leftarrow}(\tilde{t}_n) \leq t &\Leftrightarrow \tilde{t}_n \leq f(t) \\ &\Leftrightarrow \tilde{N}(f(t)) \geq n \\ &\Leftrightarrow N(t) \geq n . \end{aligned}$$

This characterizes the sequence $\{t_n\}$, thus we obtain that $t_n = f^{\leftarrow}(\tilde{t}_n)$. The assumptions on f imply the following properties of $f^{\leftarrow}$.

- The jumps of $f^{\leftarrow}$ correspond to the intervals (s_n, s_{n+1}) where f is constant. More precisely, if f is constant on (s_n, s_{n+1}), then $f^{\leftarrow}$ has a jump at $f(s_n)$ of size $s_{n+1} - s_n$. Since $f^{\leftarrow}$ is left continuous, it holds that

  $$f^{\leftarrow}(f(s_n)) = s_n , \quad \lim_{u\to f(s_n), u>f(s_n)} = s_{n+1} .$$

  Thus the jumps of $f^{\leftarrow}$ are of size C_0 at most.

- If f is increasing on an interval (s_n, s_{n+1}), then $f^{\leftarrow}$ is differentiable on $(f(s_n), f(s_n^-))$ and $(f^{\leftarrow})'(t) \leq \delta_1^{-1}$ for all $t \in (f(s_n), f(s_n^-))$.

- The jumps of f create no singularity in $f^{\leftarrow}$. If $f(s_n) > f(s_n^-)$, then $f^{\leftarrow}$ is constant on the interval $(f(s_n^-), f(s_n))$.

Let $\lceil x \rceil$ denote the smallest integer greater than or equal to the real number x. Then, for $0 \leq s \leq t$,

$$0 \leq f^{\leftarrow}(t) - f^{\leftarrow}(s) \leq C_0 \left\lceil \frac{t-s}{\delta_0} \right\rceil + \delta_1^{-1}(t-s) .$$

Thus, there exits constants c_1, c_2 such that for all $s \leq t$,

$$0 \leq f(t) - f(s) \leq c_1 + c_2(t-s) .$$

Consider now the forward recurrence time of the point process N. Then

$$\begin{aligned} 0 \leq t_{N(s)+1} - s &= f^{\leftarrow}(\tilde{t}_{\tilde{N}(s)+1}) - f^{\leftarrow}(f(s)) + f^{\leftarrow}(f(s)) - s \\ &\leq f^{\leftarrow}(\tilde{t}_{\tilde{N}(f(s))+1}) - f^{\leftarrow}(f(s)) \leq c_1 + c_2\{\tilde{t}_{\tilde{N}(f(s))+1} - f(s)\} . \end{aligned}$$

Thus, there exists constants c_3 and c_4 such that

$$\sup_{s \geq 0} \mathbb{E}[(t_{N(s)+1} - s)^p] \leq c_3 + c_4 \sup_{s \geq 0} \mathbb{E}[(\tilde{t}_{\tilde{N}(s)+1} - s)^p]$$

$\square$

Lemma 4.8. *Let $\{\epsilon_k\}$ be a sequence of i.i.d. positive random variables with finite mean μ_ϵ. Let $\{Y_k\}$ be a stationary standard Gaussian process such that*

$$\operatorname{cov}(Y_0, Y_k) = \ell(n) n^{2H-2} \tag{4.18}$$

for $H \in (1/2, 1)$ and ℓ a slowly varying function. For $k \geq 1$, define

$$\tau_k = \epsilon_k \mathrm{e}^{\sigma Y_k} \ .$$

Then the sequence $\{\tau_k\}$ is ergodic and Assumption 2.1 holds with $\lambda^{-1} = \mu_\epsilon \mathrm{e}^{\sigma^2/2}$. If $\mathbb{P}(\epsilon_1 > 0) = 1$ the Assumption 2.2 holds with $\mu = \lambda = \mu_\epsilon^{-1} \mathrm{e}^{-\sigma^2/2}$. If moreover $\mathbb{E}[\epsilon_1^q] < \infty$ for all $q \geq 1$, then (3.3) and (3.4) hold.

Remark 4.1. If instead of (4.18) we assume that

$$\sum_{k=1}^{\infty} |\operatorname{cov}(Y_0, Y_k)| < \infty \ ,$$

then the moment requirement can be relaxed to $\mathbb{E}[\epsilon_1^3] < \infty$ to obtain (3.3) and $\mathbb{E}[\epsilon_1^5] < \infty$ to obtain (3.4).

Proof of Lemma 4.8. Note first that $\mathbb{E}[\tau_k^p] < \infty$ as long as $\mathbb{E}[\epsilon_1^p] < \infty$. By Lemma 4.6, in order to check condition (3.3), we must only prove that the induced point process has finite intensiy, i.e. there exists $t > 0$ such that $\mathbb{E}[N(t)] < \infty$. See Baccelli and Brémaud (2003, Section 1.3.5). Note that

$$\mathbb{E}[N(x)] = \sum_{k=1}^{\infty} \mathbb{P}(N(x) \geq k) = \sum_{k=1}^{\infty} \mathbb{P}(t_k \leq x) \ .$$

Thus, it suffices to prove that the series on the righthand side is summable. Denote $\mu = \mathbb{E}[\tau_k]$ and $\rho_n = \operatorname{cov}(Y_0, Y_n)$. Applying Deo et al. (2009, Proposition 1), we have

$$\mathbb{E}\left[\left|\sum_{k=1}^{n} \tau_k - n\mu\right|^p\right] = O(v_n^p)$$

with $v_n = n^H \ell(n)$. If $\mathbb{E}[\epsilon_1^p] < \infty$ for p such that $p(1-H) > 1$, for n such that $n\mu > x$, it holds that

$$\mathbb{P}(t_k \leq x) = O(x^{-1} v_k^p)$$

and this series is summable. $\square$

Lemma 4.9. *Assume that $\{\tau_k\}$ and $\{\xi_k\}$ are mutually independent stationary sequences such that $\mathbb{E}[\xi_k] = 0$, $\mathbb{E}[\tau_k^2] < \infty$ and $\mathbb{E}[\xi_k^2] < \infty$. Assume that the sequence of durations is weakly stationary and that $\operatorname{cov}(\tau_0, \tau_n) = 0(n^{-\delta})$ for some $\delta > 0$ and $\sup_{s\geq 0} \mathbb{E}[t_{N(s)+1} - s] < \infty$. Assume that $\operatorname{cov}(\xi_1, \xi_n) \sim cn^{2H-2}$, with $H \in (1/2, 1)$ and $c > 0$, and that*

$$n^{-H} \sum_{k=1}^{[n\cdot]} \xi_k \Rightarrow c' B_H$$

for some $c' > 0$. Then

$$n^{-H} \int_0^{Tt} \xi_{N(s)} \, \mathrm{d}s \Rightarrow c'' B_H(t)$$

for some $c'' > 0$.

Proof of Lemma 4.9. Denote $\mathbb{E}[\tau_k] = \mu > 0$.

$$\begin{aligned}
\int_0^T \xi_{N(s)} \, \mathrm{d}s &= \sum_{k=0}^{N(T)} \tau_{k+1}\xi_k - (t_{N(T)+1} - T)\xi_{N(T)+1} \\
&= \sum_{k=0}^{N(T)} (\tau_{k+1} - \mu)\xi_k + \mu \sum_{k=0}^{N(T)} \xi_k - (t_{N(T)+1} - T)\xi_{N(T)+1} \,.
\end{aligned}$$

By independence of $\{\tau_k\}$ and $\{\xi_k\}$, we have (assuming without loss of generality that $2H - \delta > 1$),

$$\operatorname{var}\left(\sum_{k=0}^{n} (\tau_{k+1} - \mu)\xi_k\right) = O(n^{2H-\delta}) \,.$$

Thus, $n^{-H} \sum_{k=0}^{[n\cdot]} (\tau_{k+1} - \mu)\xi_k \Rightarrow 0$. Hence by the continuous mapping theorem, it also holds that $n^{-H} \sum_{k=0}^{N(T\cdot)} (\tau_{k+1} - \mu)\xi_k \Rightarrow 0$. By independence and by assumption, $(t_{N(t)+1} - T)\xi_{N(T)} = O_P(1)$. By the continuous mapping theorem, $n^{-H} \sum_{k=0}^{N(Tt)} \xi_k \Rightarrow c' B_H(\mu^{-1}t)$. □

Lemma 4.10. *Let $\{\tau_k\}$, $\{V_k\}$ and $\{\zeta_k\}$ be sequences of random variables such that*

- *$\{\zeta_k\}$ is an i.i.d. sequence of zero-mean and unit variance random variables; $\{\tau_k\}$ and $\{V_k\}$ are sequences of positive random variables;*
- *the sequences $\{(\tau_k, V_k)\}$ and $\{\zeta_k\}$ are mutually independent;*
- *there exists $s > 0$ such that $n^{-1} \sum_{k=1}^{n} \tau_{k+1}^2 V_k^2 \xrightarrow{P} s^2$;*
- *$\sup_{k\geq 0} \mathbb{E}[\tau_{k+1}^{2+\varepsilon} V_k^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$;*
- *$\sup_{s\geq 0} \mathbb{E}[t_{N(s)+1} - s] < \infty$.*

Define $\xi_k = \zeta_k V_k$*. Then* $T^{-1/2}\int_0^{T\cdot} \xi_{N(s)}\,\mathrm{d}s \Rightarrow cB$ *for some* $c > 0$*.*

Proof. Let $\mathcal{F}_k$ be the sigma-field generated by random variables $\{\tau_{j+1}, \zeta_j, V_j, j \le k\}$. Then $\mathbb{E}[\xi_k \tau_{k+1} \mid \mathcal{F}_{k-1}] = \tau_{k+1} V_k \mathbb{E}[\zeta_k] = 0$. Thus, $\{\tau_{k+1}\xi_k\}$ is a martingale difference sequence. Under the stated assumptions, the martingale invariance principle Hall and Heyde (1980, Theorem 4.1) yields that $n^{-1/2}\sum_{k=1}^{[n\cdot]} \tau_{k+1}\xi_k \Rightarrow cB$ for some $c > 0$. As in the proof of Lemma 4.9, denote $\mathbb{E}[\tau_k] = \mu > 0$ and write

$$\int_0^T \xi_{N(s)}\,\mathrm{d}s = \sum_{k=0}^{N(T)} \tau_{k+1}\xi_k + (t_{N(T)+1} - T)\xi_{N(T)} \; .$$

By the continuous mapping theorem, we have that $T^{-1/2}\sum_{k=1}^{N(T\cdot)} \tau_k \xi_{-1} \Rightarrow \lambda c B$. As previously, the last term is a negligible edge effect. This concludes the proof. $\square$

References

Miguel A. Arcones. Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability*, 22(4):2242–2274, 1994.

Alexander Aue, István Berkes, and Lajos Horváth. Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli*, 12(4):583–608, 2006.

François Baccelli and Pierre Brémaud. *Elements of queueing theory*, volume 26 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2003.

Ole E. Barndorff-Nielsen, Peter Reinhard Hansen, Asger Lunde, and Neil Shephard. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008.

Luc Bauwens and David Veredas. The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of Econometrics*, 119(2): 381–412, 2004. Dynamic factor models.

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

Philippe Bougerol and Nico Picard. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics*, 52(1-2):115–127, 1992.

Marine Carrasco and Xiaohong Chen. Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*, 18(1):17–39, 2002.

Willa W. Chen and Clifford M. Hurvich. Estimating fractional cointegration in the presence of polynomial trends. *Journal of Econometrics*, 117(1):95–121, 2003a.

Willa W. Chen and Clifford M. Hurvich. Semiparametric estimation of multivariate fractional cointegration. *Journal of the American Statistical Association*, 98(463):629–642, 2003b.

Daryl J. Daley and David Vere-Jones. *An introduction to the theory of point processes. Vol. I: Elementary theory and methods. 2nd ed.* Probability and Its Applications. New York, NY: Springer., 2003.

Rohit Deo, Clifford Hurvich, and Yi Lu. Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics*, 131(1-2):29–58, 2006.

Rohit Deo, Clifford M. Hurvich, Philippe Soulier, and Yi Wang. Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory*, 25(3):764–792, 2009.

Rohit Deo, Mengchen Hsieh, and Clifford M. Hurvich. Long memory in intertrade durations, counts and realized volatility of nyse stocks. *Journal of Statistical Planning and Inference*, 140(12):3715 – 3733, 2010. doi: DOI:10.1016/j.jspi.2010.04.037. Special Issue in Honor of Emanuel Parzen on the Occasion of his 80th Birthday and Retirement from the Department of Statistics, Texas A&M University - Emmanuel Parzen.

Robert F. Engle and Jeffrey R. Russell. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, 66(5):1127–1162, 1998.

Eric Ghysels, Pedro Santa-Clara, and Rossen Valkanov. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131 (1-2):59–95, 2006.

Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes*. Oxford University Press, New York, third edition, 2001.

P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.

James D. Hamilton. *Time series analysis*. Princeton University Press, Princeton, NJ, 1994.

Clifford M. Hurvich and Yi Wang. A pure-jump transaction-level price model yielding cointegration, leverage, and nonsynchronous trading effects. Working Paper, Stern School of Business, New York University, 2009.

Clifford M. Hurvich and Yi Wang. A pure-jump transaction-level price model yielding cointegration. *Journal of Business & Economic Statistics*, 28(4):539–558, 2010.

Ildar Abdulovich Ibragimov and Yuri Anatolevich Rozanov. *Gaussian random processes*, volume 9 of *Applications of Mathematics*. Springer-Verlag, New York, 1978. Translated from the Russian by A. B. Aries.

Domenico Marinucci and Peter M. Robinson. Weak convergence of multivariate fractional processes. *Stochastic Processes and their Applications*, 86(1):103–120, 2000.

Jean-Luc Prigent. Option pricing with a general marked point process. *Mathematics of Operations Research*, 26(1):50–66, 2001.

Sidney I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.

Sidney I. Resnick. *Heavy-tail phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

Peter M. Robinson and Domenico Marinucci. Narrow-band analysis of nonstationary processes. *The Annals of Statistics*, 29(4):947–986, 2001.

Wim Vervaat. Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 23:245–253, 1972.

Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.